Injected carrier profiles and consequent optical mode variations in semiconductor waveguides

Cinosi, N.

Award date:
2003

Awarding institution:
University of Bath

UNIVERSITY OF BATH

Injected Carrier Profiles and Consequent Optical Mode Variations in Semiconductor Waveguides

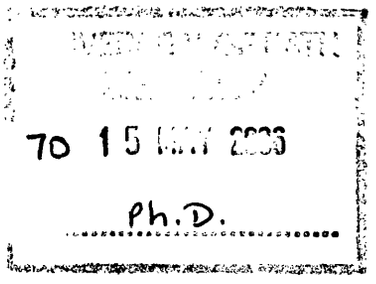
submitted by N. Cinosi

for the degree of Doctor of Philosophy

of the University of Bath

2003

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346


70 15
Ph.D.

Summary:

The aim of the thesis is to investigate the control of optical propagation in planar dielectric rib waveguides. Quasi-analytic models are suggested as accurate and fast alternative solutions to pure numerical methods with the intention of exploring the design and the feasibility of novel structures.

The analysis has been divided in two parts: the electromagnetic optical field propagation along the waveguide and the injection of carriers as a means to change the optical propagation characteristics.

The electromagnetic analysis focuses on the field propagation along longitudinally uniform and non-uniform waveguides, by using functional expansion techniques to describe the expression of the field. The Spectral Index Method, usually adopted to solve field propagation along uniform rib waveguides, is modified and adapted to the case of tapered rib waveguides and Y-junction structures.

The two-dimensional carrier distribution, subsequent to current injection inside rib waveguides, is determined by applying the Lumped Iterative Method, which reduces the two-dimensional problem to the separate solution of two coupled one-dimensional problems. Various recombination processes are considered, and included in the model in a general form, so that a large variety of materials and structures can be modelled with minor modifications.

Both optical field and carrier distribution results obtained with the methods developed in the thesis are compared with those obtained by using more detailed software packages.

Different device structures are simulated and presented, to prove the usefulness of the suggested models as a flexible computed assisted design (CAD) tool.

Acknowledgements

I deeply acknowledge my supervisor Dr. J. Sarma for his continuous creative criticism and discussions. His interest, suggestions, support and patience are integral part of this PhD work.

Beside my gratitude to Dr. F. Causa, I feel to thank all the other members, past and present, of the Optoelectronic group for generating discussions and good humour: C. Aquilanti, N. Brooks, L. Burrow, G. Busico, M. Cinalli, S. Janot, D. Masanotti, D. Mencarelli, T. Ryan, A. Tarlis, T.C. Woo, S. Yunus.

Bookham Technology, Abingdon, UK, is acknowledged for financial support. In particular Dr. I. Day and A. Harpin are thanked for discussions and let me access the software Silvaco.

I must reserve special thanks for the encouragement and the patience of my family.

Finally, I wish to strike a cheerful note for Jenny as she made my life to sound more enjoyable.

This manuscript is written with LaTeX.

Contents

List of Symbols

$\mathbf{B}$	$(Weber/m^2)$	*magnetic flux vector*
$\mathbf{D}$	(C/m^2)	*electric displacement vector*
$\mathbf{E}$	(V/m)	*electric field vector*
$\mathbf{H}$	(A/m)	*magnetic field vector*
$\mathbf{J}$	(A/m^2)	*current density vector*
$\mathbf{J_p}$	(A/m^2)	*electron current density vector*
$\mathbf{J_p}$	(A/m^2)	*hole current density vector*
k_0	(m^{-1})	*free space wavenumber*
d	(m)	*depth of the rib waveguide*
h	(m)	*height of the rib waveguide*
l	(m)	*width of the intrinsic region*
L	(m)	*length of the rib waveguide*
w	(m)	*width of the rib waveguide*
B_r	(m^3/s)	*radiative recombination coefficient*
C_n	(m^6/s)	*electron Auger recombination coefficient*
C_p	(m^6/s)	*hole Auger recombination coefficient*
e	$= 1.602 \cdot 10^{-19}\ (C)$	*electron charge*
D_n	(m^2/s)	*electron diffusion coefficient*
D_p	(m^2/s)	*hole diffusion coefficient*
n	(m^{-3})	*electron carrier concentration*
p	(m^{-3})	*hole carrier concentration*
n_0	(m^{-3})	*electron carrier concentration at equilibrium*
p_0	(m^{-3})	*hole carrier concentration at equilibrium*
N_a	(m^{-3})	*acceptor concentration in a p doped region*
N_d	(m^{-3})	*donor concentration in a n doped region*
Δn	(m^{-3})	*excess electron carrier concentration*
Δp	(m^{-3})	*excess hole carrier concentration*
G	(m^{-3}/s)	*generation rate*
R	(m^{-3}/s)	*recombination rate*
S_r	(m/s)	*surface recombination velocity*
V_t	$= 0.026\ (V)$	*thermal voltage at 300K*

β	(m^{-1})	*longitudinal propagation constant*
ϵ_0	$= 8.854 \cdot 10^{-12}\ (F/m)$	*vacuum permittivity*
ϵ	-	*dielectric constant*
η	-	*refractive index*
λ	(m)	*wavelength*
μ_0	$= 4\pi \cdot 10^{-7}\ (H/m)$	*vacuum permeability*
μ	-	*relative permeability*
μ_n	$(m^2/V-s)$	*electron mobility*
μ_n	$(m^2/V-s)$	*hole mobility*
ρ	(C/m^3)	*charge density*
σ	$(ohm^{-1}m^{-1})$	*material conductivity*
τ	(s)	*lifetime constant*
ω	(rad/s)	*angular frequency*

Chapter 1

Introduction

The continuous demand for functional integrated optoelectronic components requires the development of novel devices with complex operations. In most cases of semiconductor based devices the operational characteristics are actuated by altering the optical properties of the dielectric material.
This purpose can be technologically achieved in different ways, either during the fabrication process of the device or acting dynamically, while the device is operating. Of course, the latter is far more desirable, although more difficult to achieve, since it provides a greater flexibility.
The aim of the work presented in this thesis is to analyse the different aspects of modulating the optical propagation characteristic of semiconductor devices.
The study has been mainly focused on a specific class of devices, namely planar rib waveguides with large refractive index step (deep ridge) between the optical guiding and the surrounding cladding regions. Examples of such devices are those made on Silicon-on-Insulator (SOI) technology, [1]. The reasons for this direction of analysis have been dictated by direct industrial requirements and by the fast grow of interest, recently resumed in the area of SOI devices, [2].

The configurations of devices analysed include:

- Uniform rib waveguides and coupler
- Non-uniform (tapered) rib waveguide
- Y-junction splitter
- MultiMode Interferometer (MMI)

Although these devices have different purpose and functionality, they present some common features when modelled.
In the process of altering/modulating the optical propagation three major aspects can be highlighted:

- the optical propagation in the device before modulation
- the change/modulation of the optical properties of the device
- the optical propagation in the device after modulation

These three points have been followed to establish the workplan of the modelling process and also applied to the layout of the thesis. Each topic has been tackled and solved separately and then, finally, the solutions developed for the three different problems have been combined together to produce a model for the entire process of optical modulation. In all above cases quasi-analytic techniques have been preferred to pure numerical methods, in order to have a direct understanding of the links between the physics of the device and the parameters of the model. Moreover, quasi-analytic methods are convenient for innovation/design of novel structures since they provide an immediate insight of the solutions, in a (computationally) fast and accurate way.

Optical propagation in the device before modulation

The optical propagation in a semiconductor device has been determined by solving the wave equation for an electromagnetic propagating field. A possible approach to solve the wave equation for a complicated structure is by using a pure numerical method, e.g. the Finite Element Method, [5]. However, in the present thesis a quasi-analytic method for solving the entire two-dimensional waveguide propagation is introduced. Among several quasi-analytic methods present in literature to predict the 2-D modal propagation, [6], the Spectral Index Method (SIM) has been widely preferred, [7], since it is particularly tailored to solve waveguides with large refractive index step between the core and the cladding. Moreover, the SIM is even fast in computational speed, resulting in a very efficient method, since, conversely from many other methods, [8], both accuracy and computational speed requirements are satisfied at the same time.

The SIM, originally conceived to solve longitudinally uniform waveguides, has been extended to the case of longitudinally non-uniform waveguides. The same concepts used in the SIM have been applied to develop a method for describing the field diffraction along a tapered rib waveguide.

Change of the optical properties of the material

The optical properties of a semiconductor material can be changed by altering its refractive index, η, [4]. An effective way to achieve such changes is by injecting/depleting carriers (electrons, n, and holes, p) into/out of the material, [3], [9]. The injection of a current produces a distribution of carriers, $p(x, y)$, $n(x, y)$, across the device which will, consequently, causes a refractive index change distribution, $\eta(x, y)$. Therefore, in order to establish the injected carrier distributions, p, n, a quasi-analytic method for solving the current flow equations, in both 1-D and 2-D domains, is suggested.

Optical propagation in the device after current injection

Once the refractive index has been altered the problem is to predict the effect on the field propagation. This can be done, for instance, by solving again the wave equation, but applied to the new altered refractive index profile. Although this approach is correct it becomes tedious when the distribution of the affected refractive index is inhomogeneous. Indeed, one option is to repeat the entire computation. Alternatively, to speed up the computation the refractive index change can be regarded as a perturbation to the initial (unperturbed) structure, thus applying a perturbation analysis.

Outline of the thesis:

Following the pattern of the introduction the thesis is subdivided in three parts:

I The first part is devoted to the aspects of optical propagation inside rib waveguides. After defining the problem and reviewing some of the most relevant methods of solution, [*chapter 2*], the 2-D field propagation along a uniform, [*chapter 3*], and longitudinally non-uniform (tapered), [*chapter 4*], rib waveguide is determined.

II The second part is focused on the semiconductor aspects of the devices, describing how the perturbation of the material (refractive index change) can be achieved. Particularly, the injection of carriers into SOI semiconductor devices is investigated. Following a brief review of basic semiconductor equations, [*chapter 5*], a quasi-analytic model determining the carrier distribution in both 1-D and 2-D cases is presented, [*chapters 6-7*].

III Finally, the results obtained from the two previous parts are combined together by using the perturbation analysis. The entire analysis is applied to a class of devices, and the results presented in [*chapter 8*]. Finally, conclusions and suggestions for possible future works are drawn in [*chapter 9*].

Bibliography

[1] Tang C.K., Kewell A.K., Reed G.T., Rickman A.G., Namavar F., "Development of a Library of Low-Loss Silicon-on-Insulator Optoelectronic Devices", *IEE Proc. Optoelec.*, V.143, No.5, 1996, 312-315

[2] Soref R.A., Lorenzo J.P., "All-Silicon Active and Passive Guided-Wave Components for $\lambda = 1.3$ and $1.6\mu m$", *IEEE J. Quantum Electron.*, V. QE-22, No.6, 1986, 873-879

[3] Soref R.A., Bennett B.R., "Electrooptical Effects in Silicon", *IEEE J. Quantum Electron.*, QE-23, No.1, 1987, 123-129

[4] Tang C.K., Reed G.T., Walton A.J., Rickman A.G., "Low-Loss, Single-Mode, Optical Phase Modulator in SIMOX Material", *J. Lightwave Technol.*, V.12, No.8, 1994, 1394-1400

[5] Rahman B.M.A., Davies J.B., "Finite-Element Solution of Integrated Optical Waveguides", *J. Lightwave Technol.*, No.2, 1984, 682-688

[6] Saad M.S., "Review of Numerical Methods for the Analysis of Arbitrarily-Shaped Microwave and Optical Dielectric Waveguides", *IEEE T-MTT*, V. MTT-33, No.10, 1985, 894-899

[7] Robson P.N., Kendall P.C., "Rib Waveguide Theory by the Spectral Index Method", RSP, England, 1990

[8] Chiang K.S., "Review of Numerical and Approximate Methods for the Modal Analysis of General Optical Dielectric Waveguides", *Opt. Quant. El.*, No.26, 1994, S113-S134

[9] Bennett B.R., Soref R.A., Del Alamo J.A., "Carrier-Induced Change in Refractive Index of InP, GaAs, and InGaAsP", *IEEE J. Quantum Electron.*, V.26, No.1, 1990, 113-122

Part I

Optical Field Propagation in Waveguides

Chapter 2

Optical Propagation in Dielectric Waveguides

2.1 Maxwell's equations

The distribution of the electromagnetic field in a medium is described by Maxwell's equations, [1], introduced here for the case of time harmonic dependence, i.e. for a field of the form $\tilde{\mathbf{F}}(x,y,z,t) = \mathbf{F}(x,y,z) \cdot \exp(-i\omega t)$; ω, positive and real quantity, is the angular frequency, $\mathbf{F}$ is a complex vector:

$$\nabla \cdot \mathbf{D} = \rho \tag{2.1}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.2}$$

$$\nabla \times \mathbf{E} = -i\omega\mathbf{B} \tag{2.3}$$

$$\nabla \times \mathbf{H} = i\omega\mathbf{D} + \mathbf{J} \tag{2.4}$$

∇ is a vector differential operator depending of the spatial coordinates [Appendix A]; $\mathbf{E}$, $\mathbf{H}$, are the electric and magnetic field vectors, respectively; $\mathbf{D}$, $\mathbf{B}$, are the electric displacement and the magnetic flux vectors, respectively; ρ is the charge density inside the medium; $\mathbf{J} = \mathbf{J_s} + \mathbf{J_c}$ is the total electric current density flowing inside the medium, made of the contribution of the current source, $\mathbf{J_s}$, and the conductive current, $\mathbf{J_c}$.
In the case of linear and isotropic media, these quantities can be related as

$$\begin{aligned} \mathbf{D} &= \epsilon_0\epsilon_r\mathbf{E} \\ \mathbf{B} &= \mu_0\mu\mathbf{H} \\ \mathbf{J_c} &= \sigma\mathbf{E} \end{aligned} \tag{2.5}$$

with σ the material conductivity; μ_0 is the magnetic permeability of vacuum, μ the relative magnetic permeability; ϵ_0 is the permittivity of vacuum, ϵ_r the relative dielectric constant of the medium. The dielectric constant, ϵ, in general complex, is related to the refractive index, η, as $\epsilon = \eta^2$.

In the rest of the thesis Maxwell's eqns. (2.1)-(2.4) are solved only for the case of semiconductors and, henceforth, it is $\mu = 1$ (non ferroelectric material) and $\sigma = 0$, i.e $\mathbf{J_c} = \mathbf{0}$, (non conductive material). Moreover, it is assumed propagation in a free charge, ($\rho = 0$), free source, ($\mathbf{J_s} = 0$), medium.

2.2 The wave equation

Applying the operator $\nabla\times$, [Appendix A], to eqn. (2.3) and eqn. (2.4), and considering eqns. (2.5) yields, after some algebra, the wave equations for the electric and magnetic fields, [1]

$$\nabla^2\mathbf{E} + k_0^2\epsilon\mathbf{E} = -\nabla[\mathbf{E}\cdot\frac{\nabla\epsilon}{\epsilon}] \tag{2.6}$$

$$\nabla^2\mathbf{H} + k_0^2\epsilon\mathbf{H} = -\frac{\nabla\epsilon}{\epsilon}\times\nabla\times\mathbf{H} \tag{2.7}$$

with $k_0^2 = \omega^2\epsilon_0\mu_0 = (\frac{2\pi}{\lambda_0})^2$, and λ_0 the electromagnetic field wavelength in free space. ∇^2 is the Laplacian operator, [Appendix A]. Eqns. (2.6)-(2.7) produce six scalar equations describing the six components of the electromagnetic field, in general, all coupled together.

2.3 Scalar wave equation in Rectangular Coordinates

In a Cartesian coordinate system, (x,y,z), the gradient of the dielectric constant is expressed as $\nabla\epsilon = (\partial_x\epsilon, \partial_y\epsilon, \partial_z\epsilon)$, with the use of the notation $\partial_x = \frac{\partial}{\partial x}$, etc. Thus the spatial changes of the refractive index, $\nabla\epsilon$ in eqns. (2.6), (2.7), can be highlighted as:

1. $\partial_z\epsilon$: term representing the longitudinal variation of the refractive index, due, for example, to longitudinally tapered geometries, Fig. [2.1]. In the case of longitudinally uniform waveguide ($\partial_z\epsilon = 0$) or weakly varying structure ($\partial_z\epsilon \approx 0$) this term can be neglected.

2. $\partial_x\epsilon$, $\partial_y\epsilon$: terms representing the transverse variation of the refractive index. The first term describes, for example, the presence of a rib in a planar structure, while the second one is used to describe a multilayer vertical waveguide configuration, Fig. [2.1].

In the case of a longitudinally uniform (or weakly varying) waveguide, e.g. $\partial_z\epsilon = 0$, the vectorial wave eqn. (2.6) for the electric field can be rewritten with three different equations,

one for each component of the field, as:

$$\begin{aligned}
\nabla^2 E_x + k_0^2 \epsilon E_x &= -\partial_x [E_x \partial_x ln(\epsilon) + E_y \partial_y ln(\epsilon)] \qquad (a) \\
\nabla^2 E_y + k_0^2 \epsilon E_y &= -\partial_y [E_x \partial_x ln(\epsilon) + E_y \partial_y ln(\epsilon)] \qquad (b) \\
\nabla^2 E_z + k_0^2 \epsilon E_z &= -\partial_z [E_x \partial_x ln(\epsilon) + E_y \partial_y ln(\epsilon)] \qquad (c)
\end{aligned} \tag{2.8}$$

- the terms $\left(\partial_x [E_x \partial_x ln(\epsilon)]\right)$ in eqn. (2.8a) and $\left(\partial_y [E_y \partial_y ln(\epsilon)]\right)$ in eqn. (2.8b) account for the polarisation effect on the field solution, (the polarisation describes the spatial field orientation), [1].

- the terms $\left(\partial_x [E_y \partial_y ln(\epsilon)]\right)$ in eqn. (2.8a) and $\left(\partial_y [E_x \partial_x ln(\epsilon)]\right)$ in eqn. (2.8b) imply a vectorial nature of the solution, since they are responsible for the coupling between the different field components, [2].

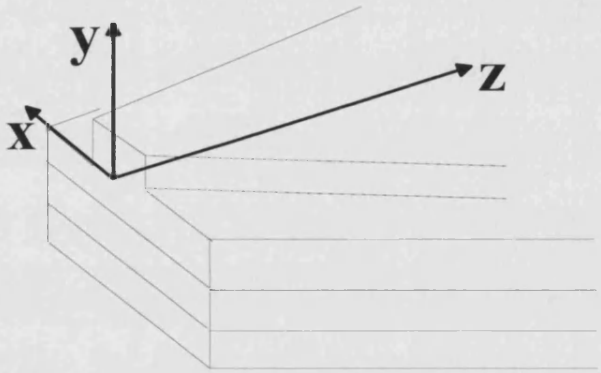


Figure 2.1: Structure of a (taper) rib waveguide in Cartesian Coordinates.

In the case of a slab waveguide (or a sequence of stratified layers), with no ribs, ($\partial_x \epsilon = 0$), the vertical component E_y is uncoupled from the other components. Indeed, eqn. (2.8b) for E_y involves only such component and is therefore an uncoupled scalar wave equation

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial y^2} + \frac{\partial^2 E_y}{\partial z^2} + k_0^2 \epsilon E_y = -\partial_y [E_y \partial_y ln(\epsilon)] \tag{2.9}$$

Moreover, if the H_y component is omitted, i.e. $H_y = 0$, it is possible to determine all the other components of the electromagnetic field from Maxwell's equations, [Appendix A]. This solution is referred to as TM_y polarisation, since the magnetic field vector lies on a plane transverse to the y axis.

In a similar manner, by starting from the wave eqn. (2.7) for the magnetic field, it is possible to write a scalar wave equation for the vertical component of the magnetic field, H_y. The other components are thus determined from Maxwell's equations, after omitting the vertical component of the electric field, i.e. $E_y = 0$. This solution is referred to as TE_y since the

electric field is transverse to the y axis.

Therefore, in a vertical slab waveguide the six components of the electromagnetic field can be entirely determined by solving two scalar Helmholtz's wave equations independently, namely for the TM_y and the TE_y polarisation:

$$\nabla^2 E_y + k_0^2 \epsilon E_y = \frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial y^2} + \frac{\partial^2 E_y}{\partial z^2} + k_0^2 \epsilon E_y = 0 \tag{2.10}$$

$$\nabla^2 H_y + k_0^2 \epsilon H_y = \frac{\partial^2 H_y}{\partial x^2} + \frac{\partial^2 H_y}{\partial y^2} + \frac{\partial^2 H_y}{\partial z^2} + k_0^2 \epsilon H_y = 0 \tag{2.11}$$

In the more general case of a rib waveguide it is $\partial_x \epsilon \neq 0$ in eqns. (2.8). Hence the presence of the rib produces coupling between TM_y and TE_y polarisation, and the field is characterised by a hybrid solution (combination of both polarisation together), [3]. A rigorous analysis will then require that all the six components of the electromagnetic field are solved together.

2.4 Scalar wave equation in Cylindrical Coordinates

Structures with cylindrical boundaries can be solved in cylindrical coordinates, (r,θ,y), Fig. [2.2]. The six electromagnetic field components in cylindrical coordinates can not be uncoupled and reduced to a scalar wave equation, even in the case of a vertical slab waveguide, i.e. if $\partial_\theta \epsilon = 0$. This is because of the form of the Laplacian operator in cylindrical coordinates, [Appendix A], [4]. However, if the field is solved specifically for the

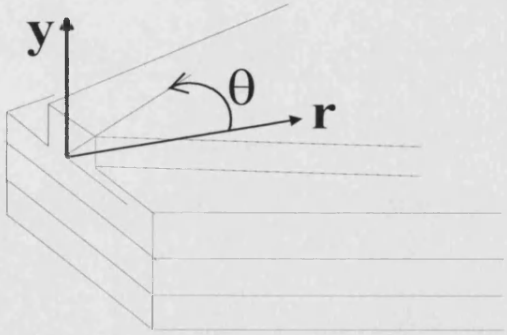


Figure 2.2: Structure of a (taper) rib waveguide in Cylindrical Coordinates.

y component, the Laplacian operator for a vector is the same as that for a scalar field, [Appendix A]. Thus eqn. (2.6), assuming a radially uniform waveguide, i.e. $(\partial_r \epsilon = 0)$,

becomes

$$\nabla^2 E_y + k_0^2 \epsilon E_y = -\partial_y [E_\theta \partial_\theta ln(\epsilon) + E_y \partial_y ln(\epsilon)] \tag{2.12}$$

Eqn. (2.12) is similar to eqn. (2.8b) and, therefore, same considerations on the presence of the rib, ($\partial_\theta \epsilon \neq 0$), as done in the previous section can be made. Similarly to the rectangular coordinate representation it is then possible to describe, in approximate terms, the electromagnetic field by solving separately the E_y and H_y components (TM_y and TE_y polarisation). In this way the main component (E_y or H_y) is determined by solving the scalar wave equation in cylindrical coordinates and the other components are obtained consequently from Maxwell's equations, [Appendix A],

$$\nabla^2 E_y + k_0^2 \epsilon E_y = \frac{\partial^2 E_y}{\partial r^2} + \frac{1}{r}\frac{\partial E_y}{\partial r} + \frac{1}{r^2}\frac{\partial^2 E_y}{\partial \theta^2} + \frac{\partial^2 E_y}{\partial y^2} + {k_0}^2 \epsilon(\theta, y) E_y = 0 \tag{2.13}$$

and similarly for H_y.

2.5 Longitudinally uniform waveguides: eigenmode problem

When the (rectangular) waveguide is uniform along the longitudinal z direction it is possible to separate the z variable from the others, expressing the field as

$$E(x, y, z) = E_m(x, y) \cdot \exp(-i\beta_m z) \tag{2.14}$$

and the scalar wave eqn. (2.10) is rewritten as

$$\frac{\partial^2 E_m}{\partial x^2} + \frac{\partial^2 E_m}{\partial y^2} + k_0^2 \epsilon(x, y) E_m = \beta_m^2 E_m \tag{2.15}$$

Eqn. (2.15) constitutes an eigenvalue equation, whose solutions, E_m, are the modes supported by the waveguide and β_m the correspondent eigenvalues describing the phase variation along the longitudinal z axis. It has been shown, [3], that the set of all possible modes supported by a uniform waveguide (i.e. guided and radiated) is complete [1].

2.6 Longitudinally non-uniform waveguides

When the (rectangular) waveguide presents a longitudinal z variation of the refractive index then the term $\partial_z \epsilon$ constitutes an extra term in eqns. (2.8), contributing to the coupling between the different components of the electromagnetic field vector. However, the approximation of weakly varying waveguide, i.e. $\partial_z \epsilon \approx 0$, reduces the full six components vectorial problem to the solution of Helmholtz's scalar eqn. (2.10), even in the case of non-uniform waveguides.

[1] By definition a set of functions $E_0, E_1, E_2, \ldots$ is **complete** in a space of functions F if every function f belonging to F can be expressed by a linear combination $f = a_0 E_0 + a_1 E_1 + a_2 E_2 + \ldots$ with $a_0, a_1, a_2, \ldots$ constant coefficients, [5].

2.6.1 Initial value problem

Although the analysis can be reduced to the scalar case, the field can not be described as in eqn. (2.14). The field propagation thus can not be reduced to an eigenmode problem, eqn. (2.15), but the field profile, $E(x, y, z)$, must be obtained by solving the more general Helmholtz's eqn. (2.10). If a pure single local mode is excited at the start of the waveguide, i.e. at $z = 0$ in Fig. [2.1], part of its power is transferred to other local guided modes and also to radiation modes. The propagation is described by the superposition of all these local modes and not only by the single mode launched initially into the waveguide.
In order to determine a unique field solution an initial distribution of the field at $z = 0$, Fig. [2.1] must be assigned. If the transverse profile of the electric field is given at $z = 0$, i.e. $E(x, y, 0) = E_0(x, y)$, then the solution of eqn. (2.10) will describe uniquely how this profile propagates and evolves along the longitudinal direction.

2.6.2 Paraxial wave equation

Although the field can not be written in separated variable form, i.e. $E(x, y, z) = f(x, y) \cdot g(z)$, for adiabatically z-varying geometries it is convenient to write

$$E(x, y, z) = F(x, y, z) \cdot \exp(-ipz) \tag{2.16}$$

The term p represents the longitudinal fast variation while $F(x, y, z)$ is a slowly varying term.
The analysis of the electromagnetic field with slowing varying pattern, $F(x, y, z)$, is greatly simplified by using the paraxial approximation expressed as

$$\left|\frac{\partial^2 F(x, y, z)}{\partial z^2}\right| \ll \left|p\frac{\partial F(x, y, z)}{\partial z}\right| \tag{2.17}$$

and

$$\left|\frac{\partial^2 F(x, y, z)}{\partial z^2}\right| \ll \left|\frac{\partial^2 F(x, y, z)}{\partial x^2}\right| \quad ; \quad \left|\frac{\partial^2 F(x, y, z)}{\partial z^2}\right| \ll \left|\frac{\partial^2 F(x, y, z)}{\partial y^2}\right| \tag{2.18}$$

Inserting eqn. (2.16) into Helmholtz's eqn. (2.10) and making use of eqns. (2.17), (2.18) the paraxial wave equation for the slowly varying term, $F(x, y, z)$, is obtained as

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - 2ip\frac{\partial}{\partial z} + k_0^2\epsilon(x, y) - p^2\right] F(x, y, z) = 0 \tag{2.19}$$

The paraxial approximation is particularly valid in the case of waveguide structures with a slow longitudinal variation of the refractive index (adiabatic geometry).

2.7 Functional expansion

For some particular waveguide configurations, like dielectric slab, uniform waveguides, etc. eqn. (2.10) can be solved analytically, while for more complicated structures the solution is found only by using numerical techniques.
One general way to represent the solution, valid for any type of waveguide, is to express the field in terms of a complete set of functions, [6]. The description of the solution is simplified if the set is chosen judiciously, and only few terms of the expansion are sufficient to reproduce the solution accurately.
The general field expansion is of the form

$$E(x,y,z) = \sum_{k=1}^{\infty} a_k(z) f_k(x,y) + \int b(s,z) r(x,y,s) \mathrm{d}s \tag{2.20}$$

$f_k(x,y)$ and $r(x,y,s)$ describe a discrete and a continuous set of functions, respectively; $a_k(z)$, $b(s,z)$ are coefficients to be determined. The union of $f_k(x,y)$ and $r(x,y,s)$ must be complete.
For an orthogonal set of basis functions the coefficients $a_k(z)$, $b(s,z)$, are related to the field through the relations

$$a_k(z) = \int_D f_k(z) E^*(x,y,z) \mathrm{d}x\mathrm{d}y \tag{2.21}$$

$$b(s,z) = \int_D r(x,y,s) E^*(x,y,z) \mathrm{d}x\mathrm{d}y \tag{2.22}$$

D, the domain of integration, is chosen as appropriate; E^* denotes complex conjugate.
The choice of the functional set depends of the waveguide geometry: example of such sets, useful for describing waveguide modes, are the Hermite-Gauss functions [7], sine/cosine functions [9], metal waveguide modes [10], etc. These sets are not only complete, but also orthogonal.
For a rigorous analysis both the contribution to the expansion (2.20) must be considered; however, in this work the continuous modes have been neglected. The concept of functional expansion will be used later in the thesis, when, in particular, the sine/cosine (Fourier) terms and the metal waveguide modes are utilised.

2.8 Review of numerical methods for waveguide propagation

When the geometry of the waveguide is complicated it may not be possible to obtain an analytic solution and some approximate method must be applied. Approximate methods can be either purely numerical or quasi-analytic. A complete review of approximate methods for field analysis of optical waveguides can be found in [11] and [12]. Amongst all of them the most common methods have been selected and reported more exhaustively

below. They have also been utilised as benchmark to test the accuracy of the results presented in this thesis.

2.8.1 Effective Dielectric Constant (EDC)

The Effective Dielectric Constant (EDC) is a method utilised to solve the two-dimensional mode propagation along a longitudinally uniform rib waveguide (eigenmode problem). Although the presence of the rib generates a hybrid field solution and it is not possible, in principle, to express the field profile, $E_m(x,y)$, in terms of separated variable functions, i.e. $E_m(x,y) \neq f(x) \cdot g(y)$, the EDC method assumes that $E_m(x,y) = f(x) \cdot g(y)$, [3].
Moreover, if the ratio between the width, w, and the height, h, of the rib, Fig. [2.3], is large and the refractive index step, $\eta_2 - \eta_1$, between the inner and the outer part of the rib is small, the variation of the field profile along the x direction is much smaller than the variation along y. It is then possible, as a first approximation, to neglect any variation along the x axis, [13], i.e. $\partial_x^2 E_m(x,y) \approx 0$, and consequently to reduce the wave eqn. (2.15) to the one-dimensional equation, with reference to Fig. [2.3],

$$\frac{\partial^2 g(y)}{\partial y^2} + [{k_0}^2 \epsilon(y) - \beta_q^2] g(y) = 0 \qquad \text{q} = \text{inner, outer} \tag{2.23}$$

Eqn. (2.23) is equivalent to the slab wave equation with known analytic solution. The slab structure is solved twice, in two different regions, corresponding to the inner and outer parts of the rib, Fig. [2.3a]. The eigenvalue propagation constants β_q are then utilised as new (effective) dielectric constants of an (effective) slab waveguide uniform along y, Fig. [2.3b]. Because of the uniformity, $\partial_y^2 E_m(x,y) = 0$, the slab wave equation to be solved is

$$\frac{\partial^2 f(x)}{\partial x^2} + [\beta_q^2 - \beta_m^2] f(x) = 0 \tag{2.24}$$

The solution of eqn. (2.24) describes the modal propagation constant, β_m, of the mth mode.

The polarisation effect is also taken into account: if the scalar electric field is directed along x, $E_x(x,y)$, then eqn. (2.23) is for a TM polarised slab, while eqn. (2.24) refers to TE polarised slab. Conversely, for an electric field directed along y, $E_y(x,y)$, the polarisation of the slab is TE for eqn. (2.23) and TM for eqn. (2.24), [3].

2.8.2 Finite Element Method (FEM)

The Finite Element Method (FEM) is one of the most accurate methods for solving optical waveguide propagation, [14]: it is often considered as the "benchmark" for comparison of solutions obtained by other methods. In the FEM the entire structure is first subdivided into a finite number of triangular/rectangular regions called elements, Fig. [2.4]. Using many elements, ij, it is possible to approximate the description of any continuum geometry,

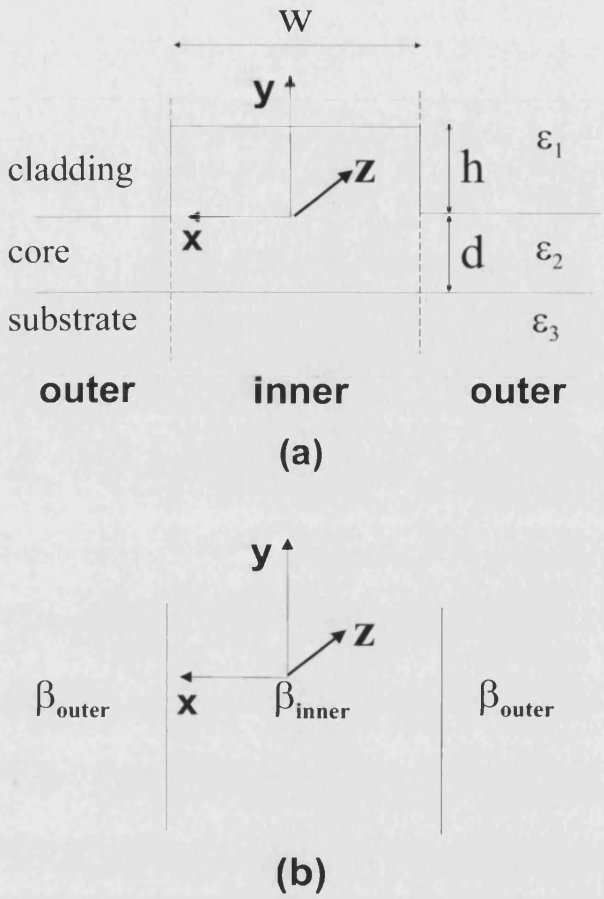


Figure 2.3: EDC method applied to a rib waveguide: (a) the inner and outer regions are solved separately as slab waveguides uniform along the x axis; (b) the propagation constants obtained in case (a) are utilised as effective dielectric constants to solve an effective slab waveguide uniform along the y axis.

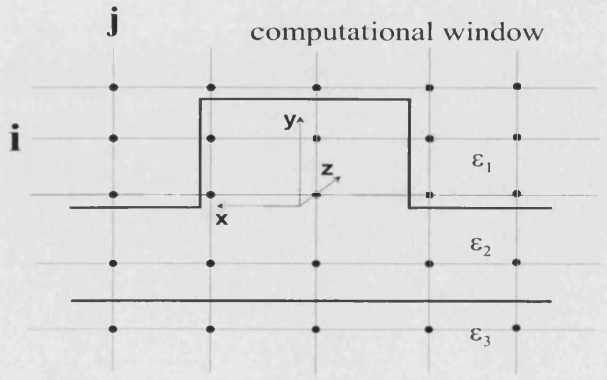


Figure 2.4: Finite Element Method grid.

with arbitrary refractive index distribution and complex boundaries. The field functions are defined by a set of polynomials over each element, E_{ij}, whose coefficients must be determined. In this way, using a variational analysis, the problem is reduced to a matrix equation, [15]. The field at each node, E_{ij}, is coupled with the field at the surrounding nodes, $E_{i\pm1,j\pm1}$ or with the boundary conditions if the node is at the boundary of the computational window. The accuracy of the solution depends mainly on the number of nodes chosen in the discretisation process and the shape of the elements. A large number of elements will approximate more precisely the physical geometry of the device, but will increase the computational complexity. However, for some geometries, many elements of the matrix are zero, therefore they do not contribute to the computation of the field, and both computer memory and computation times are reduced.
One advantage of the FEM is that it can be used for a full vectorial analysis, including all the six components of the electromagnetic field, [8]. In the case of a scalar analysis many nodes of the matrix are uncoupled and the computation process is much faster.

2.8.3 Local Mode Expansion (LME)

The Local Mode Expansion (LME) is a method for describing the field propagation along longitudinally non-uniform waveguides, e.g. tapers. It is based on a functional expansion, eqn. (2.20), with the set of functions, $f_k(x,y)$, given by the so called Local Normal Modes, [6]. At any given point z the field is expressed in terms of the modes of a hypothetical straight waveguide with same width as that of the actual waveguide at the point z, Fig. [2.5]. These local modes used for the expansion are not themselves solution of Helmholtz's eqn. (2.10) for the tapered geometry, but they can be superimposed to represent the field of the actual non-uniform waveguide, satisfying eqn. (2.10).
Since locally there is no z variation for these modes, they can be obtained by solving the eigenvalue eqn. (2.15). These modes are locally orthogonal. For a particular z the total field can be locally expressed as

$$E(x,y;z) = \sum_k a_k f_k(x,y)\exp(-i\beta_k z) + \int a(s)f(s,x,y)\exp[-i\beta(s)z]\mathrm{d}s \qquad (2.25)$$

Two different cross sections, (i) and (j), along the z direction, generate two virtual straight waveguides with width w_i and w_j, respectively, Fig. [2.5]. The two different widths produce a step discontinuity at the common interface $z = z_0$.

The continuous taper can therefore be approximated by a series of discontinuous step regions of uniform slabs. By sub-dividing the taper into a large number of small steps, and propagating the field from one section to the next section, and repeating this process at each step, the longitudinally non-uniform field propagation is obtained.
Focusing the attention to two adjacent sections, (i) and (j), assuming that an incident field, $E_i(x,y,z)$, is propagating from section (i) towards section (j), the discontinuity at $z = z_0$

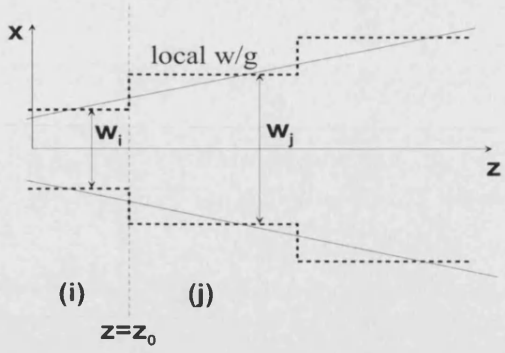


Figure 2.5: The tapered waveguide can be approximated by a cascade of uniform waveguides with different widths.

generates a reflected field, E_r, back into section (i), as well as a transmitted field, E_t, into section (j).

Each field, reflected and transmitted, is expanded using the modes of the local waveguide where they are propagating, (i) and (j), respectively. The reflected field, E_r, is:

$$E_r(x,y,z) = \sum_k a_k^i f_k^i \exp(-i\beta_k^i z) + \int a^i(s) f^i(s) \exp[-i\beta^i(s)z]\mathrm{d}s \tag{2.26}$$

f_k^i and $f^i(s)$ are the local guided and radiation modes of section (i), respectively. The transmitted field, E_t, is:

$$E_t(x,y,z) = \sum_k a_k^j f_k^j \exp(-i\beta_k^j z) + \int a^j(s) f^j(s) \exp[-i\beta^j(s)z]\mathrm{d}s \tag{2.27}$$

f_k^j and $f^j(s)$ are the local guided and radiation modes of section (j), respectively.

The expansion coefficients, a_k^i, a_k^j, $a^i(s)$, $a^j(s)$ are determined by imposing the continuity of the total electric field at the step interface $z = z_0$,

$$E_i(x,y,z_0) + E_r(x,y,z_0) = E_t(x,y,z_0) \tag{2.28}$$

and similarly for the magnetic field, $H(x,y,z)$.

If the step discontinuity is small enough it is possible to assume that the reflected field is negligible, [16], and that the amount of incident field coupled to the radiated transmitted modes is also negligible, [16], i.e.,

$$E_t(x,y,z) \approx \sum_k a_k^j f_k^j \exp(-i\beta_k^j z) \tag{2.29}$$

It is then possible to determine the coefficients a_k^j as the overlap integral

$$a_k^j = \frac{\int_D E_i(x,y,z_0) f_k^{j*}(x,y)\mathrm{d}x\mathrm{d}y}{\int_D |f_k^j|^2 \mathrm{d}x\mathrm{d}y} \tag{2.30}$$

The domain of integration D is the transverse cross-section interface between the two steps, $z = z_0$, Fig. [2.5].

2.8.4 Beam Propagation Method (BPM)

The Beam Propagation Method (BPM) is a method for determining the field diffraction in a half-space with arbitrary inhomogenities: it predicts the response of a given structure to an input field. Although the presence of inhomogenities requires the coexistence of both forward and reverse propagating waves the BPM considers only the forward travelling waves. Similarly to the Local Mode Expansion the BPM progresses stepwise, assuming that the dielectric profile remains constant within a section. Inside each section the BPM propagates the field as diffracting in free space and treats the inhomogenities as a perturbation of the free space propagation. This procedure is valid under the condition of weakly varying waveguide, and it formally solves the paraxial wave equation, [17].
A qualitative physical interpretation of the BPM is schematised in Fig. [2.6]. For each section the BPM procedures are:

- the input field is Fourier transformed into plane waves
- each plane wave is diffracted (propagation in uniform medium) by half section
- the plane wave spectrum is recomposed into the field profile
- the phase profile is corrected by the presence of the inhomogenities
- the field is Fourier transformed again into plane waves
- each plane wave is propagated in the remaining half section
- the plane wave spectrum is recomposed into the field profile which is used as input field for the next section

These steps are repeated at each successive section of the waveguide.
The sensitiveness of the algorithm to the length of the sections depends on the refractive

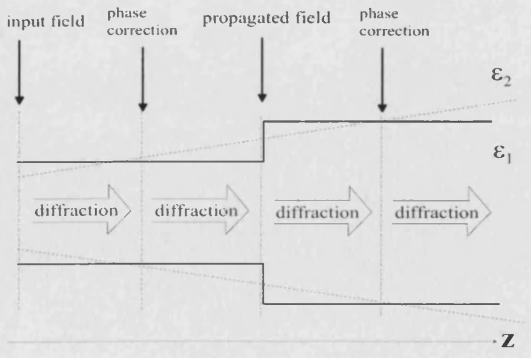


Figure 2.6: Beam propagation along a tapered device.

index step. It has been shown, [18]-[19], that for slowly varying taper waveguides the BPM works well and provides accurate results.

Bibliography

[1] Ramo S., Whinnery J.R, Van Duzer T, "Field and Waves in Communication Electronics", John Wiley & Sons, NY

[2] Ghatak A.K., Thyagarajan K., "Optical Electronics", CUP, Cambridge, 1989

[3] Rozzi T., Mongiardo M., "Open Electromagnetic Waveguides", IEE, London, 1997

[4] Stinson D.C., "Intermediate Mathematics of Electreomagnetics", Prentice-Hall, London, 1976

[5] Kreyszig E., "Advanced Engineering Mathematics", John Wiley & Sons, New York, 1993

[6] Marcuse D., "Theory of Dielectric Optical Waveguides", Academic Press, NY, 1974

[7] Causa F., "Hermite-Gauss Functions in the Analysis of a Category of Semiconductor Optical Devices", *PhD thesis*, Univ. of Bath, 1998

[8] Rahman B.M.A., Davies J.B., "Vector-H Finite Element Solution of GaAs/GaAlAs rib waveguides", *IEE Proc., part J*, V.132, No.6, 1985, 349-353

[9] Henry C.H., Shani Y., "Analysis of Mode Propagation in Optical Waveguide Devices by Fourier Expansion", *IEEE J. Quantum Electron.*, V.27, No.3, 1991, 523-530

[10] Collin R.E., Foundations for Microwave Engineering, McGraw-Hill, Tokyo, 1966

[11] Chiang K.S., "Review of Numerical and Approximate Methods for the Modal Analysis of General Optical Dielectric Waveguides", *Opt. Quant. El.*, No.26, 1994, S113-S134

[12] Robertson M.J., Ritchie S., Dayan P., "Semiconductor Waveguides: Analysis of Optical Propagation in Single Rib Structures and Directional Coupler", *IEE Proc., part J*, V.132, No.6, 1985, 336-342

[13] Payne F.P., "A new Theory of Rectangular Optical Waveguides", *Opt. Quant. El.*, QE No.14, 1982, 525-537

[14] Rahman B.M.A., Davies J.B., "Finite-Element Solution of Integrated Optical Waveguides", *J. Lightwave Technol.*, No.2, 1984, 682-688

[15] Rahman B.M.A., Davies J.B., "Finite-Element Analysis of Optical and Microwave Waveguides Problems", *IEEE Trans. on MTT*, 32, 1984, 20-28

[16] TsuTsumi K., Imada Y., Hirai H., Yuba Y., "Analysis of Single-Mode Optical Y-Junctions by the Bounded Step and Bend Approximation", *J. Lightwave Technol.*, V.6, No.4, 1999, 590-600

[17] März R., "Integrated Optics. Design and Modeling", Artech House, London, 1995

[18] Sewell P., Benson T.M., Kendall P.C., Anada T., "Tapered Beam Propagation", *Electron. Lett.*, V.32, No.11, 1996, 1025-1026

[19] Sujiecki S., Sewell P., Benson T.M., Kendall P.C., "Novel Beam Propagation Algorithms for Tapered Optical Structures", *J. Lightwave Technol.*, V.17, No.11, 1996, 2379-2388

Chapter 3

Analysis of 2-dimensional uniform waveguides: the Spectral Index Method

The Spectral Index Method (SIM) is a quasi-analytic method, [1], suitable for solving the field propagation along longitudinally uniform rib waveguides (eigenmode problem) characterised by a large refractive index step between the core guiding region and the top cladding region, Fig. [3.1]. The SIM is based on the functional expansion of the field in terms of two suitable sets of functions: the rectangular metal waveguide modes and the Fourier spectrum. The method can be summarised in three steps:

- the electric field inside the rib, region A in Fig. [3.1], is expanded in terms of metal waveguide modes
- the electric field underneath the rib, region B in Fig. [3.1], is decomposed in plane waves (Fourier transformed along the x axis)
- the above two expressions for the field are matched at the interface between the two regions, A and B, i.e. the base of the rib, $y = 0$, utilising a variational technique

As a result, the matching conditions generate an equation (dispersion relation) whose numerical solutions yield the propagation constants, β, of the modes supported by the waveguide. The value of β is then used in the expansion expressions to obtain the electric field profile. Note that for the region B under the rib the field is expressed in the spectral domain, thus Inverse Fourier transform from spectral to spatial domain yields the field profile.

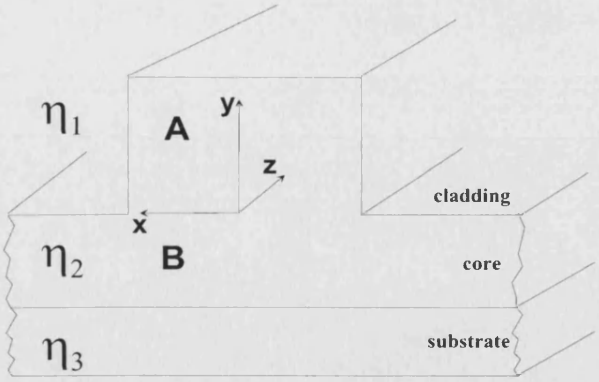


Figure 3.1: Structure of a rib waveguide.

3.1 Evanescent boundary conditions and Effective Rib

As introduced in Section 2.3 the electromagnetic field propagating in a rib waveguide is characterised by a hybrid solution. The two polarisation TE_y and TM_y are coupled together and all the six field components are necessary to describe the propagation. However, it has been shown, [1], that if the rib is surrounded by a cladding material with a much lower refractive index the coupling between the two polarisation is negligible. This feature is particularly true when the core/cladding interface is straight. In the case of a rib waveguide, Fig. [3.1], the top wall and the two side walls are both straight, therefore, away from the corners and the shoulders of the rib the above assumption is valid. In proximity of corners and shoulders there are sudden changes of the interface; however, since for guided modes the field is mainly confined across the central region away from the corners, it is acceptable to neglect the coupling effect due to the corners.

A consequence of the uncoupling between TE_y and TM_y polarisation is that the field can still be described accurately by using only five components, [2], and, more importantly, the problem can be reduced to the solution of the scalar Helmholtz's equation, eqn. (2.10). Because the waveguide is longitudinally uniform the field is expressed as in eqn. (2.14), each $E_m(x, y)$ representing a different mode of the waveguide, and β_m the correspondent propagation constant. Henceforth the index m is omitted for simplicity. The scalar Helmholtz equation is:

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + k_0^2\epsilon(x, y) - \beta^2\right]E(x, y) = 0 \tag{3.1}$$

Eqn. (3.1) is solved separately for the vertical component of the magnetic, H_y, and electric, E_y, field in order to obtain the TE_y, TM_y modes, respectively. However, in [1] it has also been proved that in the case of TE_y polarisation, solving the scalar wave equation for the component H_y is equivalent to solving the horizontal component of the electric field, E_x. Therefore, the field for the TE_y polarisation can be determined by treating E_x as the main component, Tab. [3.1].

field components		TE_y		TM_y
electric		$\mathbf{E_x}$, E_z		E_x, $\mathbf{E_y}$, E_z
magnetic		H_x, H_y, H_z		H_x, H_z

Table 3.1: Definition of TE_y, TM_y polarisation field components using SIM. In bold are the main components that are solved in the scalar Helmholtz's equation.

The large refractive index step between the core and the cladding, i.e. $\eta_1 << \eta_3 < \eta_2$, Fig. [3.1], produces guided (bound) modes that are mostly confined inside the core region. As a consequence, the field in the outer cladding region decays rapidly after a short distance from the core/cladding surface, and its value is almost negligible away from the interface.

The decaying behaviour is described by evanescent boundary conditions which replace the physical boundary conditions at the core/cladding boundary, [Appendix B],

$$\frac{\partial E_n}{\partial n} = -\frac{E_n}{d_n} \qquad \frac{\partial E_t}{\partial n} = -\frac{E_t}{d_t} \tag{3.2}$$

E_n and E_t are the normal and tangential components of the electric field to the boundary surface, respectively; ∂_n is the derivative along the direction normal to the surface; the distances d_n and d_t describe the penetration depth of the field inside the cladding layer before its amplitude becomes negligible. These distances are determined as, [Appendix B]:

$$d_t = \frac{1}{\sqrt{\beta^2 - k_0^2 \epsilon_1}} \tag{3.3}$$

$$d_n = \frac{\epsilon_1}{\epsilon_2} \frac{1}{\sqrt{\beta^2 - k_0^2 \epsilon_1}} \tag{3.4}$$

β is the propagation constant of the field as described in eqn. (3.1). It is important to note that because β is the sought parameter, the evanescent boundary conditions (3.2) can not be defined until the final solution is known: this would invalidate the use of such boundary conditions. However, in the case of guided modes, the propagation constant β must be in the range $k_0\eta_1 << k_0\eta_3 < \beta < k_0\eta_2$. It is thus reasonable to approximate $\beta \approx k_0\epsilon_2$ in order to evaluate the distances d_n, d_t, and define the boundary conditions to the problem. This approximation has negligible effect on the final value of β, [1].

A detailed explanation and proof of eqns. (3.2)-(3.4) is given in Appendix B.

Because the distances d_n, d_t, are, in general, small, the field is almost zero after a short distance away from the rib. It is then convenient to replace the actual rib with an *Effective Rib*, Fig. [3.2], whose field outside is assumed to be perfectly zero, [3].

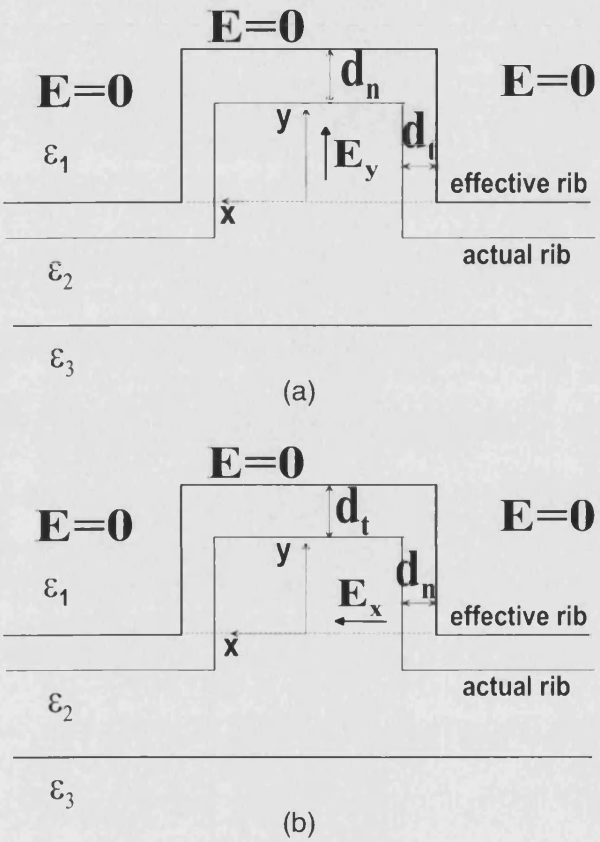


Figure 3.2: Replacement of the rib waveguide with an effective rib: (a) TM_y polarisation; (b) TE_y polarisation.

3.2 Modal expansion inside the rib

With the new defined *effective rib* the wave eqn. (3.1) is applied first to the region $y > 0$, omitting, for the moment, the layers below the base of the rib. The electric field inside region A, Fig. [3.3], is expressed by using a functional expansion (Section 2.7). The artifice

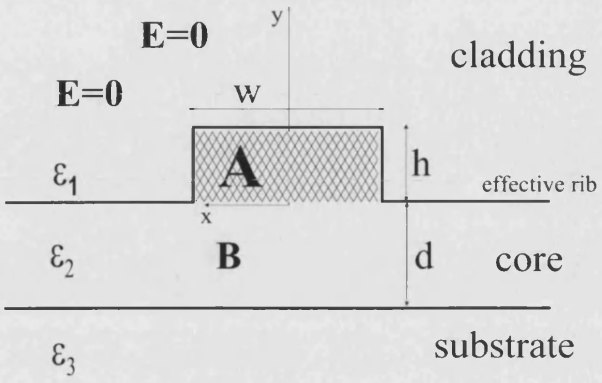


Figure 3.3: Region A: inside the rib.

of the effective rib, surrounded outside by zero electric field, introduces new boundary conditions similar to those applied for parallel metal plates. The two vertical walls of the effective rib represent in fact two metal plates. It then seems expedient to expand the field along the x direction in terms of metal waveguide modes (complete set of functions); thus,

$$E_A(x,y) = \sum_{k=1}^{\infty} F_k(x) \cdot G_k(y) \tag{3.5}$$

where the function $F_k(x)$ is the kth order mode of two metal parallel plates separated by a distance w, Fig. [3.3],

$$\begin{aligned} F_k(x) &= \cos\left[\tfrac{(2k-1)\pi}{w}x\right] && \text{for symmetric modes} \\ F_k(x) &= \sin\left[\tfrac{2k\pi}{w}x\right] && \text{for anti-symmetric modes} \end{aligned} \tag{3.6}$$

In [1] it has also been shown that for the kth eigenmode of the rib waveguide only the kth term of the expansion (3.5) contributes significantly to the expression of the field, while the other terms are negligible. Therefore, each mode can be solved by using only one appropriate term of the expansion; for example, the first (symmetric) mode is described in region A as, ($k = 1$):

$$E_A(x,y) \approx \cos(s_A x) \cdot G(y) \qquad s_A = \frac{\pi}{w} \tag{3.7}$$

Substituting this expression into eqn. (3.1) it is:

$$\frac{\mathrm{d}^2 G(y)}{\mathrm{d}y^2} + [{k_0}^2 \epsilon_2 - s_A^2 - \beta^2] G(y) = 0 \tag{3.8}$$

The new Helmholtz's eqn. (3.8) is equivalent to a slab waveguide problem uniform in x. The expression for $G(y)$ is given analytically as

$$G(y) = A_1 \sin(k_{2A} y) + A_2 \cos(k_{2A} y) \tag{3.9}$$

with

$$k_{2A}^2 = k_0^2 \epsilon_2 - s_A^2 - \beta^2 \tag{3.10}$$

The constants A_1, A_2 are determined by applying the boundary conditions at the top and bottom of region A. At the top of the rib, $y = h$, the electric field is zero, due to the effective rib. The second condition should be assigned at the base of the rib, $y = 0$. Because the region B below the rib is solved by using a different expansion to represent the electric field, this condition will be imposed explicitely later when the two expressions for the electric field are matched at $y = 0$.

3.3 Modal expansion underneath the rib

Similarly to region A, also in region B, underneath the rib, Fig. [3.4], the field is expressed through a functional expansion. In region B there are no variations along the x axis, and the waveguide is supposed to extend between $x = -\infty$ and $x = +\infty$. Conversely, along the vertical y direction the waveguide shows a sequence of layers, each one characterised by a different refractive index. The bottom layer, extending ideally to $y = -\infty$, represents the substrate of the device, while the base of the rib and the core/cladding interface at $y = 0$ constitutes the top of the sequence.

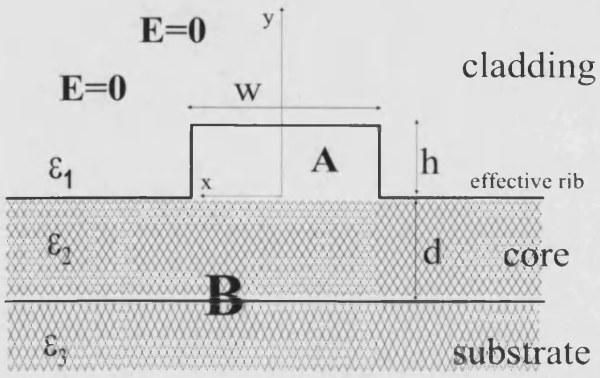


Figure 3.4: Region B: underneath the rib.

Because of the uniformity, in each layer the field can be Fourier expanded along the x direction

$$E_B(x,y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathcal{E}_B(s,y) \exp(-isx) \mathrm{d}s \tag{3.11}$$

$\mathcal{E}(s,y)$ is the Fourier transform of the electric field $E_B(x,y)$ in the x-axis

$$\mathcal{E}_B(s,y) = \int_{-\infty}^{+\infty} E_B(x,y) \exp(isx) \mathrm{d}x \tag{3.12}$$

Inserting eqn. (3.11) into Helmholtz's eqn. (3.1) yields, for each vertical layer,

$$\frac{\partial^2 \mathcal{E}_B(s,y)}{\partial y^2} + [k_0^2 \epsilon(y) - s^2 - \beta^2]\mathcal{E}_B(s,y) = 0 \tag{3.13}$$

The dielectric constants ϵ is $\epsilon = \epsilon_2$ in the core layer, $-d < y < 0$, and $\epsilon = \epsilon_3$ in the substrate, $y < -d$. Eqn. (3.13) is equivalent to the slab wave equation with an effective dielectric constant, ϵ_s,

$$\epsilon_s = \epsilon - \frac{s^2}{{k_0}^2} \tag{3.14}$$

The field underneath the rib is thus solved in the Fourier Transform domain, in terms of the spectral wavenumber s. This reduces the problem from a two-dimensional rib waveguide to a one-dimensional slab. The analytic solutions of eqn. (3.13) are:

$$\mathcal{E}_B(s,y) = \begin{cases} B_1 \sin(k_{2B}y) + B_2 \cos(k_{2B}y) & -d < y < 0 \quad \text{core} \\ C_1 \exp(k_{3B}y) & y < -d \quad \text{substrate} \end{cases} \tag{3.15}$$

with k_{2B} and k_{3B} vertical wavenumbers inside the core and the substrate region, respectively,

$$\begin{aligned} k_{2B}^2 &= k_0^2\epsilon_2 - s^2 - \beta^2 \\ k_{3B}^2 &= \beta^2 + s^2 - k_0^2\epsilon_3 \end{aligned} \tag{3.16}$$

The constants B_1, B_2, C_1, are determined by applying the interface conditions. The type of solution reflects the fact that the field is bound inside the core region, while it is decaying in the substrate layer, thus, assuming an exponential decaying type of solution for the substrate region, $y < -d$. It is important to note that, in contrast to the region A, the electric field is now not expressed as the product of two functions with separate variables. The x variable is linked to the y variable through the spectral variable s, since the field in the spectral domain, $\mathcal{E}_B(s,y)$, depends of both the s and the y variables. The general two-dimensional physical nature of the problem due to the presence of the rib is still preserved, although the method of solution has been simplified.

3.4 Interface conditions

In order to determine $\mathcal{E}_B(s,y)$ uniquely, both the continuity of the electric field and its derivative are applied at the interface between the core and the substrate regions, $y = -d$. The continuity of the field can be equivalently applied to its Fourier Transform in the spectral domain, thus the interface conditions at $y = -d$ are:

$$\mathcal{E}_B(s,-d^-) = b\mathcal{E}_B(s,-d^+) \quad (a) \qquad \frac{\partial \mathcal{E}_B(s,-d^-)}{\partial y} = \frac{\partial \mathcal{E}_B(s,-d^+)}{\partial y} \quad (b) \tag{3.17}$$

with $b = 1$ for the TE_y polarisation (E_x tangential to the interface) or $b = \epsilon_2/\epsilon_3$ in the TM_y case (E_y perpendicular to the interface).

The continuity of the electric field is also required at the bottom of the rib, $y = 0$, between the two regions A and B, except that the field at the two sides of this interface is represented by two different expansions. In the top part, inside the rib, the field as given by eqn. (3.7) is

$$E_A(x, y = 0) = \cos(s_A x) \cdot G(y = 0) \tag{3.18}$$

Underneath the rib the field is Fourier expanded, thus

$$\mathcal{E}_B(s, y = 0) = \int_{-\infty}^{+\infty} E_B(x, y = 0) \exp(isx) \mathrm{d}x \tag{3.19}$$

In order to match the profile it is necessary to write both the expressions in the same domain. For convenience the matching is constrained in the spectral domain s

$$\int_{-w/2}^{+w/2} \cos(s_A x) \cdot G(y = 0) \exp(isx) \mathrm{d}x = \mathcal{E}_B(s, y = 0) \tag{3.20}$$

Before imposing the continuity of the derivative of the electric field it must be pointed out that because in region A the expression of the field has separated variables solution, while the same is not true in the bottom region B, *it is impossible to achieve the exact matching of the derivative of the electric field*, [1]. This is due to the fact that the field inside the rib has been represented by using only one term of expansion (3.5), and the approximated expression (3.7) does not satisfy exactly the wave equation.
However, a second condition is still needed in order to determine all the coefficients B_1, B_2, C_1 uniquely in eqn. (3.15). The best option is to minimise the mismatching between the two derivatives of the field at the two sides of the common interface, $y = 0$. This condition is realised by using a variational technique and allowing a small discontinuity for the derivative, $\partial E/\partial y$, but achieving the stationarity for the propagation constant β, [4]. This particular choice is justified since the propagation constant β is the sought parameter. It is, therefore, more desirable to minimise the error introduced in β rather than the error generated in the field profile, $E(x, y)$. In fact, the error introduced in eqn. (3.1) by the approximate expression (3.7) can be reflected either on the field profile, $E(x, y)$ or in the value of the propagation constant, β.
The application of the interface conditions (3.17) and (3.20), together with the variational method, provide an equation (dispersion relation) whose solutions are the propagation constants, β. The variational technique and the consequent dispersion relation are described in Appendix C.

To evaluate the anti-symmetric modes of the rib waveguide it is necessary to begin with the anti-symmetric metal waveguide modes in the expansion (3.5). As an example, for the first anti-symmetric mode eqn. (3.7) is replaced by

$$E_A(x, y) = \sin(s_A x) \cdot G(y) \qquad s_A = \frac{2\pi}{w} \tag{3.21}$$

The Fourier expansion in region B, underneath the rib, remains unchanged, since it is general and accounts for any type of field.
Similarly, higher order modes, either symmetric or anti-symmetric, can be found by utilising the correspondent metal mode, eqn. (3.6), for the expression of the field in region A. Alternatively, they can be found by searching for the solutions β of higher order in the dispersion relation determined for the fundamental mode.

3.5 Computed results: single rib waveguide

The parameters for the waveguide analysed are described in Tab. [3.2], with reference to Fig. [3.3]. Three different cladding materials are modelled, namely air ($\eta_1 = 1$), Silicon

$\lambda_0 = 1.55\mu m$		
$w = 3.52\mu m$	$d = 2.64\mu m$	$h = 1.63\mu m$
$\eta_2 = 3.4764$	$\eta_3 = 1.447$	

Table 3.2: Parameters for the uniform rib waveguide analysed.

Dioxide ($\eta_1 = 1.447$) and Nitride N_2 ($\eta_1 = 2$).
Tab. [3.3] summarises the effective propagation constants, $\eta_{ff} = \beta/k_0$, obtained by the SIM, for both TE_y and TM_y polarisation. These results are compared with those obtained by applying the Effective Dielectric Constant method (EDC) and the Finite Element Method (FEM).
The FEM method has been implemented by using the commercial software Silvaco installed at the company Bookham Technology. Silvaco provides a high-level user interface to choose the meshing of the structure. The user can select amongst three default meshing options: coarse, medium and fine. At high-level usage, the user does not access directly to the meshing parameters. But the software Silvaco, for each of the three options, generates a grid, by using fixed set up meshing parameters (the meshing procedure is not adaptive). For the present case the finest meshing option has been chosen.
Fig. [3.5] shows the contour plot of a typical field amplitude profile, $E(x, y)$, for the first symmetric mode.
According to Tab. [3.3] both SIM and EDC provide accurate results comparable with the FEM, which is considered the benchmark, although the SIM reaches a better accuracy. However, there are some situations beyond the limits of the applicability of the EDC, when the use of the SIM becomes relevant and necessary. As stated in Section 2.8.1, the EDC is valid for shallow rib waveguides, when the ratio w/h is large. Moving away from this condition the results obtained by the EDC will deviate from those given by SIM. This effect is shown in Fig. [3.6].

	EDC, η_{ff}	SIM, η_{ff}	FEM, η_{ff}
TE_y			
Air	3.469554	3.468094	3.468184
SiO_2	3.469567	3.468128	3.468217
N_2	3.469593	3.468201	3.468318
TM_y			
Air	3.469150	3.467784	3.467874
SiO_2	3.469174	3.467823	3.467973
N_2	3.469222	3.467905	3.468049

Table 3.3: Effective propagation constants $\eta_{ff} = \beta/k_0$ (dimensionless) obtained by using SIM. Results are compared with those obtained by the EDC and FEM (benchmark). Results for three different cladding materials (air, SiO_2, N_2) are reported, for both TE_y and TM_y polarisation.

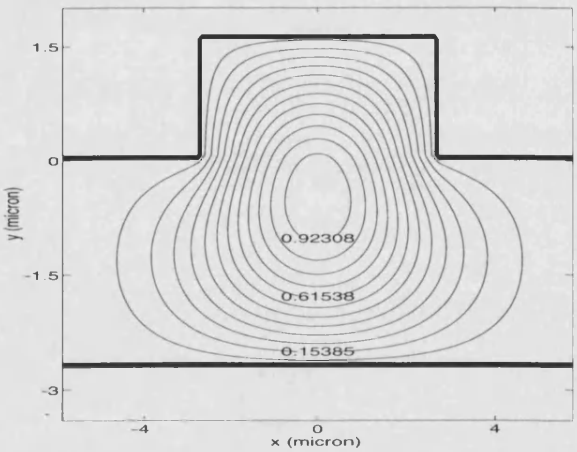


Figure 3.5: Rib waveguide: SIM computed contour plot of the field amplitude profile for the fundamental mode, in Arbitrary Units (A.U.).

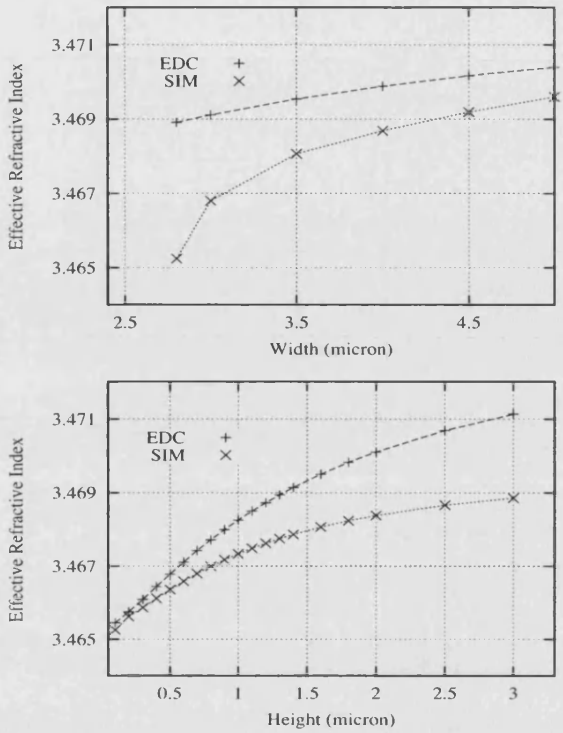


Figure 3.6: Comparison of the effective propagation constant, $\eta_{ff} = \beta/k_0$ (dimensionless), obtained by applying the EDC and the SIM: (a) the width of the rib, w, is varied, (b) the height of the rib, h, is varied.

3.5.1 Waveguide coupler

Another important application where the use of the SIM is relevant since the EDC becomes inaccurate is the waveguide coupler, [5]-[7]. A two ribs waveguide coupler, with center-to-center distance of $T = 2.48\mu m$, Fig. [3.7], and same parameters as for the rib above, is analysed.

The Spectral Index Method can be easily extended to the coupler configuration by replacing the expression (3.6) of the field inside the single rib with

$$F_k(x) = \begin{cases} A_1 \cos\left[\frac{(2k-1)\pi}{w}(x+T/2)\right] & (-w-T)/2 < x < (w-T)/2 \\ A_1 \cos\left[\frac{(2k-1)\pi}{w}(x-T/2)\right] & (T-w)/2 < x < (w+T)/2 \qquad \text{for symmetric modes} \\ 0 & elsewhere \end{cases}$$

$$F_k(x) = \begin{cases} A_1 \cos\left[\frac{(2k-1)\pi}{w}(x+T/2)\right] & (-w-T)/2 < x < (w-T)/2 \\ A_1 \sin\left[\frac{2k\pi}{w}(x-T/2)\right] & (T-w)/2 < x < (w+T)/2 \qquad \text{for anti-symmetric modes} \\ 0 & elsewhere \end{cases} \tag{3.22}$$

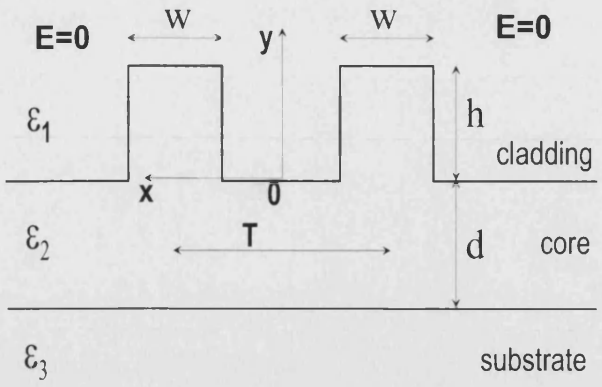


Figure 3.7: Structure of a two ribs waveguide coupler.

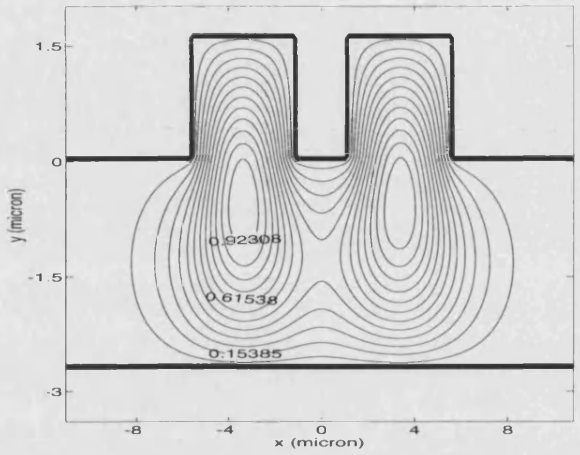
The field representation in the region below the ribs remains the same as in eqn. (3.11), i.e. Fourier transformed along x, because of uniformity.

Results obtained by using SIM, EDC and FEM are compared in Tab. [3.4], for TE_y polarisation and different cladding materials. The waveguide coupler is characterised by two mode solutions, namely the symmetric and the anti-symmetric modes, with correspondent propagation constants, β_{sym} and β_{asym}. Typical field amplitude profiles for the symmetric and the anti-symmetric modes of a coupler are shown in Fig. [3.8].

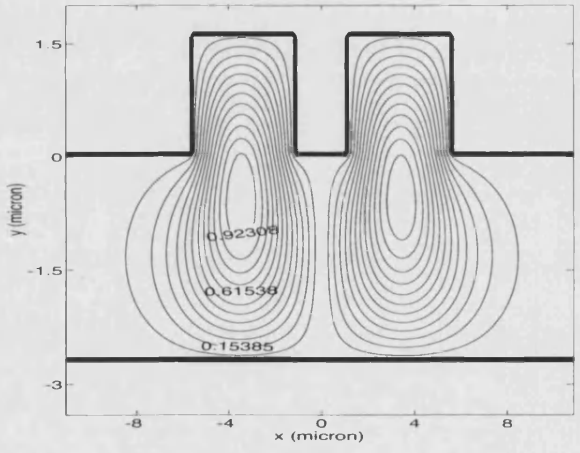
If an input field is launched into one of the two ribs, this field will gradually couple to the secon rib. After a certain distance, L (coupling length), the field is entirely transferred to the second rib, [6],

$$L = \frac{\pi}{(\beta_{sym} - \beta_{asym})} \tag{3.23}$$

Although both EDC and SIM provide values of β which are sufficiently accurate and comparable with the FEM, the EDC becomes inaccurate for the estimation of the coupling length. Indeed, to obtain accurate difference between the two propagation constants, $(\beta_{sim} - \beta_{asym})$, is particularly important rather than the accuracy of the values of the propagation constants themselves. This feature can be better appreciated by referring to Tab. [3.4], and taking the FEM results as the "benchmark". If EDC produces β which is in error by 0.03% from the benchmark, it leads to an error in the range of 20% in the coupling length. The SIM produces a far better accuracy in η_{ff} which translates to a typical error of 0.02% in the coupling length.



(a)



(b)

Figure 3.8: Coupler waveguide: SIM computed contour plot of field amplitude profile, in Arbitray Units (A.U.): (a) symmetric mode, (b) anti-symmetric mode.

	EDC, η_{ff}	SIM, η_{ff}	FEM, η_{ff}
symmetric mode: β_{sym}/k_0			
Air	3.469805	3.468381	3.468506
Si/SiO_2	3.469805	3.468309	3.468403
N_2	3.469805	3.468275	3.468370
anti-symmetric mode: β_{asym}/k_0			
Air	3.469307	3.467978	3.468100
Si/SiO_2	3.469307	3.467907	3.467999
N_2	3.469307	3.467872	3.467964
coupling length (μm), L			
Air	1559	1922	1910
SiO_2	1559	1924	1920
N_2	1559	1924	1911

Table 3.4: Effective propagation constants β/k_0 and coupling length of a coupler waveguide, for different cladding material.

Bibliography

[1] Robson P.N., Kendall P.C., "Rib Waveguide Theory by the Spectral Index Method", RSP, England, 1990

[2] Stern M.S., Kendall P.C., McIlroy P.W.A., "Analysis of the Spectral Index Method for Vector Modes of Rib Waveguides", *IEE Proc., Part J*, V.137, No.1, 1990, 21-26

[3] McIlroy P.W.A., Stern M.S., Kendall P.C., "Spectral Index Method for Polarised Modes in Semiconductor Rib Waveguides", *J. Lightwave Technol.*, V.8, No.1, 1990, 113-117

[4] Kendall P.C., Stern M.S., Burke S.V., CPlanar "Waveguide Analysis by the Spectral Index Method. I: Rib and Uniformly Buried Waveguides", *Opt. Quant. El.*, No.25, 1993, 771-787

[5] Ng W.-C., Stern M.S., "Analysis of Multiple-Rib Waveguide Structures by the Discrete Spectral Index Method", *IEE Proc., Optoelectron.*, V.145, No.6, 1998, 365-371

[6] Marcuse D., "Theory of Dielectric Optical Waveguides", Academic Press, NY, 1974

[7] Ng W.-C., Chua S.-J., "The Design of a Twin-Rib Waveguide Optical Switch by the Discrete Spectral Index Method", *Microw. and Optic. Tech. Lett.*, V.18, No. 5, 1998, 337-338

Chapter 4

Field propagation in tapered deep ridge waveguides

As stated in Chapter 2, in the case of longitudinally non-uniform waveguides the electric field can not be written as in eqn. (2.14) and therefore the propagation can not be described as an eigenmode problem. It is then necessary to solve the more general Helmholtz's eqn. (2.10) to determine how a transverse field profile, given at a specific initial point, propagates along the structure. In this chapter an approximate method is suggested for solving a two-dimensional taper waveguide by using the concepts of the Spectral Index Method adapted to the present geometry.

4.1 Taper deep ridge waveguide

The analysis is focused on evaluating optical (electromagnetic) wave propagation in linearly varying rib waveguide (taper), Fig. [4.1]. It is also assumed that the refractive index step between the core and the surrounding cladding is large (deep ridge).

For this structure the width of the tapered rib changes linearly along the (longitudinal) z

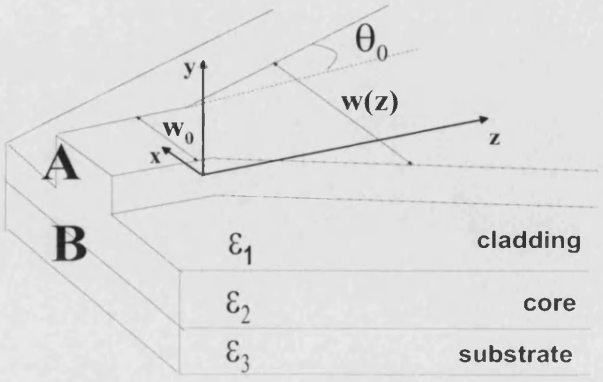


Figure 4.1: Structure of a linearly tapered rib waveguide.

direction, varying from an initial aperture, w_0, at one edge of the taper, to a value of $w(z)$, at any distance z:

$$w(z) = w_0 + 2(\tan\theta_0)z \tag{4.1}$$

θ_0 is the half-angle of the taper.

Because of the large refractive index step between the core and the cladding the taper rib, in region A, is replaced by an effective (tapered) rib surrounded outside by zero electric field, Fig. [4.2]. The extra dimensions d_n, d_t, of the effective rib are the same as those introduced in Section 3.1, eqns. (3.3), (3.4).

Similarly to the uniform waveguide case, in Chapter 3, the replacement of the actual rib

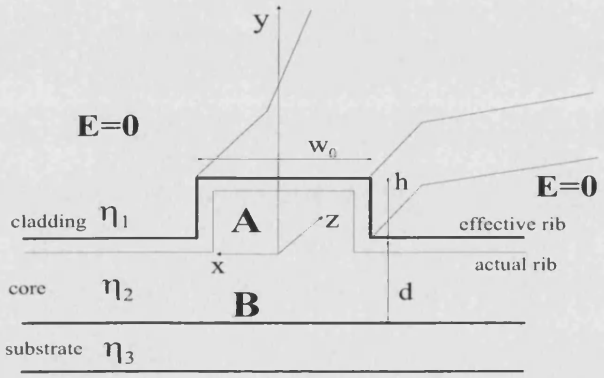


Figure 4.2: Linearly tapered effective rib.

with an effective rib introduces evanescent boundary conditions. Therefore, for weakly varying tapers, i.e. $\partial_z \epsilon \approx 0$, it is possible to uncouple the transverse electromagnetic field components. The vectorial field analysis is thus reduced to the solution of the scalar Helmholtz's eqn. (2.10) applied to the E_y and H_y components separately, for TM_y and TE_y polarisation, respectively, as described in Tab. [3.1].

Moreover, in the case of small angle θ_0 (slowly varying taper) it is convenient to separate the slow and the fast variation of the field expression, as introduced in Section 2.6.2, and solve the paraxial wave equation (2.19) rather than Helmholtz's equation (2.10):

$$E(x, y, z) = F(x, y, z) \cdot \exp(-ipz) \tag{4.2}$$

with $F(x, y, z)$ the slowly varying term and p the fast propagating term. Henceforth in the present chapter the analysis is directed to tapers with small flare angle θ_0 (adiabatic), which satisfy the paraxial conditions (2.17)-(2.18).

Similarly to the SIM, the analysis of the waveguide is sub-divided in to two regions, namely inside the rib (region A), and underneath the rib (region B), Fig. [4.1]. Following the same approach as the SIM, the field in region B is Fourier Transformed along the x axis, while a convenient functional form is specified for the field inside the rib. Since this functional expression determines and "drives" the spatial distribution of the field in the bottom region

B, the choice of the expression of the field inside the rib must be judicious.
As initial field a bound mode of the uniform input rib waveguide (connected to the taper section) is launched into the taper, Fig. [4.1]. Because the input waveguide is longitudinally uniform it supports mode solutions and it is reasonable to analyse specifically how each of these modes propagate along the tapered section.

4.2 Modal expansion inside the rib

Inside the rib (region A), Fig. [4.1], the slowly varying function, $F(x,y,z)$, is expressed by using the same metal-like modes, eqn. (3.6), as for the longitudinally uniform case, but including the variation of the rib width explicitly inside the functional form of such modes. Thus, assuming that the argument of the cosine functions varies along z, accordingly to the variation of the taper width, $w(z)$, the electric field, inside region A, is expressed as

$$F_A(x,y,z) = \sum_{k=1}^{\infty} A_k(y,z) \cdot \cos\left[\frac{(2k-1)\pi}{w(z)}x\right] \tag{4.3}$$

The basis set of functions

$$\cos\left[\frac{(2k-1)\pi}{w(z)}x\right] \qquad k = 1,2,\ldots \tag{4.4}$$

is not complete. However, since it is expected that the field profile varies similarly to such functions, the assumption (4.3) seems physically reasonable. The expression (4.3) thus constitutes an "educated" approximation of the field, which satisfies in approximate terms the wave equation for the given taper geometry.
The longitudinal, z, and lateral, x, directions are linked explicitly through the terms (4.4) to include the effects of the taper, but the longitudinal dependence is retained even in the coefficients $A_k(y,z)$, in order to maintain the field representation as general as possible.
As a first approximation it is considered only one term of expansion (4.3), similarly to what assumed in the SIM for uniform waveguides, i.e.,

$$F_A(x,y,z) = H(x,z) \cdot G(y) = a(z) \cdot \cos\left[\frac{\pi}{w(z)}x\right] \cdot G(y) \tag{4.5}$$

Since only one specific bound mode of the input waveguide is launched into the taper the choice of only one term in eqn. (4.3), i.e. the term corresponding to such mode, is sufficient to describe the field. In fact, the adiabatic property of the taper does not alter significantly the shape of the propagating field profile from that of the initial field.
It must be emphasised that the cosine term alone in eqn. (4.5) is not sufficient to describe the entire nature of the propagation because, then, the total power transmitted across the cross-section at any z will not be the same as that transmitted across any other cross-section at a different value of z; thus the term $a(z)$ is necessary in eqn. (4.5) to account for

longitudinal power conservation.

4.3 The slowly varying term $a(z)$

By inserting expression (4.5) into the paraxial wave equation (2.19) it is possible to determine a functional form for the slowly varying term $a(z)$ as

$$a(z) = \frac{1}{\sqrt{w(z)}} \exp\Big(-i\frac{q}{p}z\Big) \qquad q = q_R - iq_I \tag{4.6}$$

with p the fast propagating term and q a complex constant. The entire derivation of eqn. (4.6) is given in Appendix D.
An initial arbitrary but reasonable value for the fast propagating term p is chosen and, consequently, the values of the slowly varying term $a(z)$ are determined.
q_R describes the phase correction to the arbitrarily chosen fast propagation p, so that the total phase variation $p + q_R/p$ satisfies the taper geometry.
The amplitude term of $a(z)$:

$$\frac{\exp(-\frac{q_I}{p}z)}{\sqrt{w(z)}} = \frac{\exp(-\frac{q_I}{p}z)}{\sqrt{w_0 + 2(\tan\theta_0)z}} \tag{4.7}$$

accounts for the field amplitude decrease as it propagates longitudinally. The form (4.7) is in accordance with the fact that a diffracting field decreases as $\sim \frac{1}{\sqrt{z}}$. This specific behaviour complies with the principle of energy conservation over the cross-section at any z.

4.4 Modal expansion underneath the rib

Below the rib, $y < 0$, region B in Fig. [4.1], the structure of the waveguide is the same as the longitudinally uniform waveguide. A more general expression for the field can be used and it is not necessary to specify explicitly the dependence between x and z directions, as done for the rib part in eqn. (4.5). Any effect due to the presence of the taper in the top region will be enforced by the functional form (4.5) and by satisfying the continuity conditions at the common interface, $y = 0$. The term $a(z)$ is maintained explicitly only for convenience, together with the fast propagating term, p, as follows,

$$E_B(x,y,z) = F_B(x,y,z) \cdot \exp(-ipz) = a(z) \cdot L(x,y,z) \cdot \exp(-ipz) \tag{4.8}$$

Similarly to the analysis of a longitudinal uniform waveguide, here too the field is Fourier Transformed along the x axis and solved in the spectral domain, s,

$$\mathcal{F}_B(s,y,z) = \int_{-\infty}^{+\infty} F_B(s,y,z) \cdot \exp(isx)\mathrm{d}x = a(z) \cdot \mathcal{L}(s,y,z) \tag{4.9}$$

with $\mathcal{L}(s,y,z)$ the Fourier Transform of $L(x,y,z)$ with respect to x.
Making use of eqn. (4.9) the paraxial wave eqn. (2.19) becomes

$$\frac{\partial^2\mathcal{L}(s,y,z)}{\partial y^2}\cdot a(z)+[k_0^2\epsilon-s^2-p^2]\cdot\mathcal{L}(s,y,z)\cdot a(z)=2ip\frac{\mathrm{d}a(z)}{\mathrm{d}z}\cdot\mathcal{L}(s,y,z)+2ip\frac{\partial\mathcal{L}(s,y,z)}{\partial z}\cdot a(z) \tag{4.10}$$

The two terms on the right-hand-side of eqn. (4.10) account for the effects of the taper. Of these two terms the second is neglected, i.e.,

$$\frac{\partial\mathcal{L}(s,y,z)}{\partial z}=\frac{\partial L(s,y,z)}{\partial z}\approx 0 \tag{4.11}$$

This assumption can be justified because both the amplitude and the (small) phase perturbation due to the taper are described by the term $a(z)$ in the top region, where the taper occurs, and there are no other non-uniformities beneath the rib which generates longitudinal field variation in excess to that produced by the taper rib. Thus any field variation occurring in the bottom part is due to and determined by the top part.
A physical interpretation of eqn. (4.10) can be given by regarding the left-hand side (LHS) terms as a longitudinally uniform waveguide equation similar to eqn. (3.13). The right hand side (RHS) can be regarded as a perturbation term to the uniform propagation. Thus, by assuming that $\frac{\partial\mathcal{L}}{\partial z}\approx 0$ the perturbation effect is dictated only by the presence of the taper in the top part.
The term $a(z)$ and its derivative in eqn. (4.10) are estimated by making use of the expression (4.6). Therefore it is possible to determine an analytic form for the term $\mathcal{L}(s,y,z)$ as

$$\mathcal{L}(s,y,z)=\begin{cases} B_1(z)\sin(k_{2B}y)+B_2(z)\cos(k_{2B}y) & -d<y<0 \quad \text{core} \\ C_1(z)\exp(k_{3B}y) & y<-d \quad \text{substrate}\end{cases} \tag{4.12}$$

with k_{2B} and k_{3B} given by:

$$\begin{aligned} k_{2B}^2&=k_0^2\epsilon_2-s^2-p^2+4ip\tfrac{\tan\theta_0}{w(z)}-q \\ k_{3B}^2&=p^2+s^2-k_0^2\epsilon_3-4ip\tfrac{\tan\theta_0}{w(z)}+q\end{aligned} \tag{4.13}$$

and B_1, B_2, C_1 coefficients to be determined by applying the boundary conditions at the common interface $y=-d$.

4.5 Interface conditions

The solution of the entire field is uniquely determined by enforcing the interface conditions at the common planes $y=0$ and $y=-d$, Fig. [4.2]. Similarly to the SIM at $y=-d$ it must be

$$\mathcal{F}_B(s,-d^-,z)=b\mathcal{F}_B(s,-d^+,z)\qquad\frac{\partial\mathcal{F}_B(s,-d^-,z)}{\partial y}=\frac{\partial\mathcal{F}_B(s,-d^+,z)}{\partial y} \tag{4.14}$$

with $b = 1$ for TE_y and $b = \epsilon_2/\epsilon_3$ for TM_y polarisation.
The continuity of the field at the bottom of the rib region, $y = 0$, Fig. [4.2], is enforced in the spectral domain

$$\int_{-w/2}^{w/2} F_A(x,y,z) \cdot \exp(isx)\mathrm{d}x \equiv \int_{-w/2}^{w/2} H(x,z) \cdot G(y=0) \cdot \exp(isx)\mathrm{d}x = \mathcal{F}_B(s, y=0, z) \tag{4.15}$$

Differently from the uniform waveguide the conditions (4.14) and (4.15) are z dependent and yet must be satisfied at any value of z. This requirement is accomplished in exact terms since it is automatically built in the method. In fact, the same way as these conditions were satisfied for each value of the spectral variable s in the case of the uniform waveguide, eqns. (3.17), (3.20), now they are satisfied for each value of the spectral variable s and the longitudinal variable z.
As final condition the continuity of the field derivative is required at the bottom of the rib. Of course, since the field inside the rib has been approximated with an expression that does not satisfy exactly the wave equation it will not be possible to accomplish such condition in exact terms. However this limitation is overcome by applying a variational technique.

4.6 Variational analysis

The fast propagating term p, eqn. (4.2), is fixed in order to assign, as initial condition, a mode of the input waveguide. Therefore, the only term undetermined in the field expression, $E(x,y,z)$, is the term $a(z)$. The variational technique is applied by adjusting the slowly varying term $a(z)$ with the intent of minimising the field derivative mismatch. As derived in Appendix C, the variational form for a tapered device is given by, eqn. (C.18),

$$\int_{-\infty}^{+\infty}\int_{z_0}^{+\infty} \left[F_A^* \frac{\partial F_A(x, y=0^+, z)}{\partial y} - F_B^* \frac{\partial F_B(x, y=0^-, z)}{\partial y}\right]\mathrm{d}x\mathrm{d}z + \frac{\iiint_V 2ipF_{A,B}^* \partial_z F_{A,B}\mathrm{d}x\mathrm{d}y\mathrm{d}z}{\iiint_V F_{A,B}F_{A,B}^*\mathrm{d}x\mathrm{d}y\mathrm{d}z} = 0 \tag{4.16}$$

The first integral in eqn. (4.16) is similar to the variational expression for the longitudinally uniform rib waveguide, eqn. (C.6), [3], and accounts for the field derivative mismatch at the base of the rib. The second integral is due to the presence of the taper and describes the fraction of energy that is lost by the initial mode as it propagates along the taper (i.e. coupling to other guided or radiated modes). Since the expression (4.5) chosen for the field is not enough to describe any other modes, but the mode launched at the beginning of the taper, the second integral is essential to represent (as a global effect) such other modes (guided and radiated), so that the total transmitted energy is conserved at the transverse cross-section, for any z, [2].

4.7 Choice of the fast propagating term p

The value of the fast propagating term p in eqn. (4.6) must be chosen much larger than the phase correction, i.e. $p \gg q_R/p$, so that $a(z)$ satisfies the slowly varying condition. In order to assign as initial field a mode of the uniform input rib waveguide it is necessary to chose p as the value of the propagation constant β of such mode. In fact, at $z = 0$ the phase correction term q_R/p vanishes and, if $p = \beta$ the expression of the field inside the rib region, eqn. (4.2), is equivalent to that of the uniform waveguide, eqn. (3.7).

4.8 Results: optical field distribution

The model suggested here in Chapter 4 is applied to deep ridge adiabatic taper waveguides with different flare angle; the dimensions and the refractive indices are summarised in Tab. [4.1].

The fundamental mode of the straight input rib waveguide is launched as initial field along

$\lambda_0 = 1.55\mu m$		
$w = 3.52\mu m$	$d = 2.64\mu m$	$h = 1.63\mu m$
$\eta_1 = 1$ (Air)	$\eta_2 = 3.4764$	$\eta_3 = 1.447$

Table 4.1: Parameters for the tapered rib waveguide analysed.

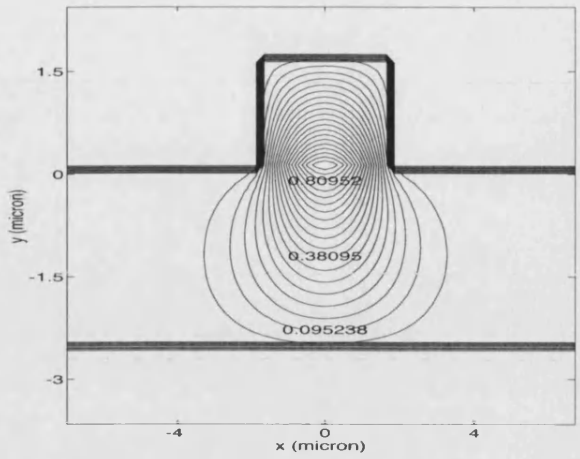
the tapered device. The propagation constant of such a mode is obtained by applying the SIM, Tab. [3.3].

Figs. [4.3b-c] show how the field $E_0(x, y)$ launched at $z = 0$, Fig. [4.3a], spreads out as it propagates along the taper. However, further along z, as the waist of the taper becomes larger the ridge tends to confine the field more in the lateral x direction. The increase of the ridge width in effect increases the horizontal confinement of the field (same as in slab waveguides), considering that both the etch depth and refractive indices are kept constant.

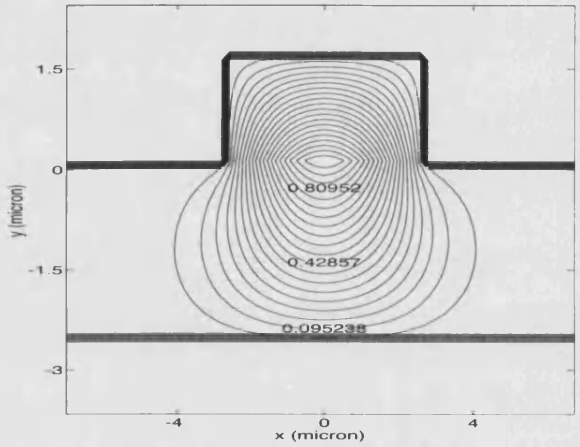
The field in region A is always entirely confined inside the ridge, as enforced by the assumption of external zero electric field. The total field confinement inside the top region and the gradual broadening of the field in the region underneath can also be observed in the longitudinal contour plots, Fig. [4.4].

4.8.1 Comparison with BPM and LME methods

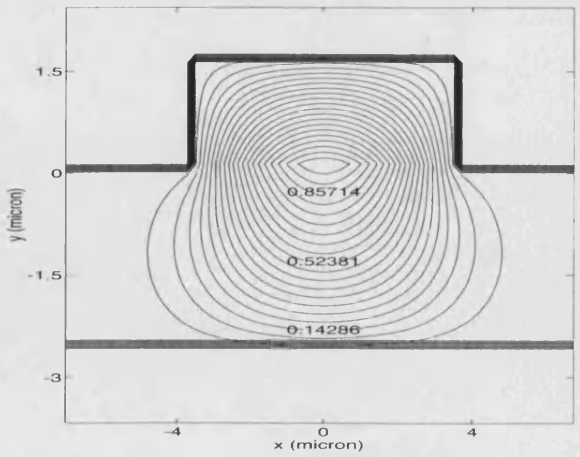
In order to test the accuracy of the present model results are compared with those obtained by Yamauchi et al., [5], by applying a Finite-Difference BPM algorithm (Section 2.8.4). The taper structure analysed has a flare angle of $\theta_0 = 0.5^o$ and the field is propagated for a



(a)

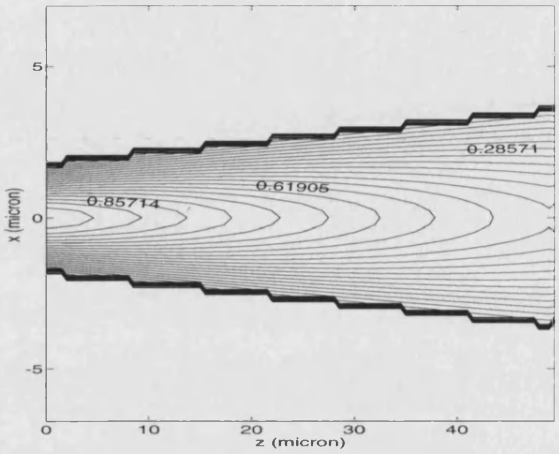


(b)

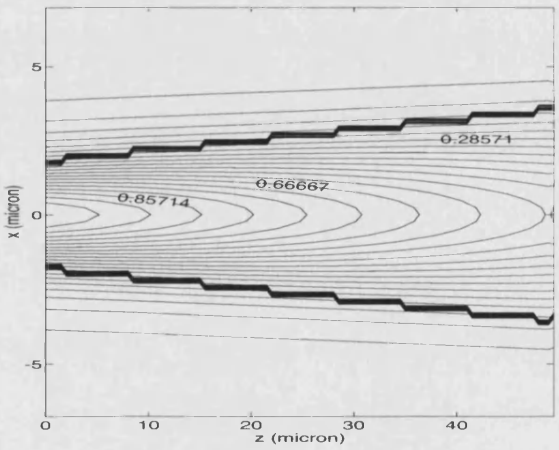


(c)

Figure 4.3: Field profile contour plot, in Arbitrary Units (A.U.), at different transverse cross sections, Fig. [4.1]: (a) $z = 0$, initial field fed by the input rib waveguide; (b) $z = 50\mu m$; (c) $z = 100\mu m$. Angle $\theta_0 = 2^o$.



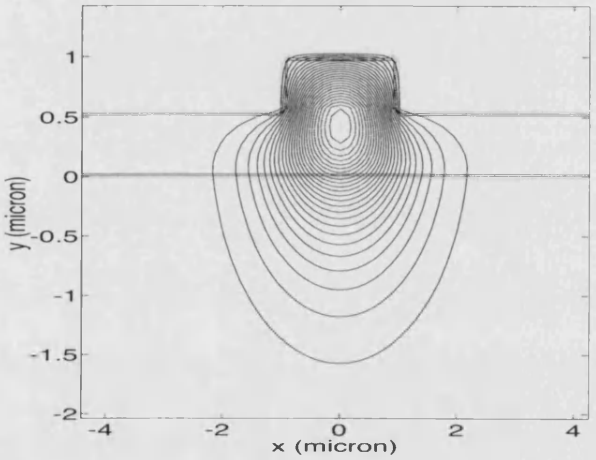
(a)



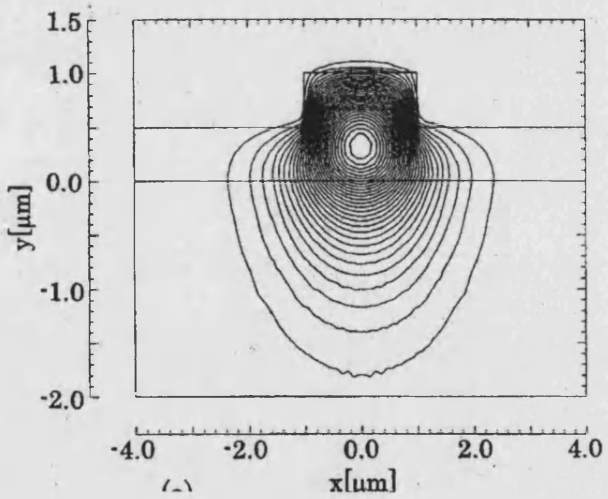
(b)

Figure 4.4: Field profile contour plot, in Arbitrary Units (A.U.), at different longitudinal cross sections, Fig. [4.1]: (a) top rib region A, $y = 0.5\mu m$; (b) bottom region B, $y = -0.5\mu m$. Angle $\theta_0 = 2^o$.

distance of $50\mu m$; the material refractive indices are: $\eta_2 = 3.44$, $\eta_3 = 3.40$. Because of the small refractive index step between the core and the substrate there is less field confinement in the vertical direction, and the field spreads more into the substrate layer.
Fig. [4.5] shows that field profiles at the output of the guide obtained by the two different methods are similar. In both cases the adiabatic taper retains the single peaked pear shape of the field, centred at the same location in the middle of the core region. The two profiles are normalised in amplitude, so that the lateral spreading at the two sides of the rib can be compared.



(a)



(b)

Figure 4.5: Comparison of results with BPM. Field profile contour plot, in Arbitrary Units (A.U.), at the output transverse section of the taper. (a) present model; (b) Beam Propagation Method. Angle $\theta_0 = 0.5^o$.

Results are also compared with those obtained by using the Local Mode Expansion techniques (Section 2.8.3). Fig. [4.6] compares the field longitudinal contour plot in the bottom region. Both methods predict the increasing lateral confinement of the field as it propagates along the longitudinal direction, z.

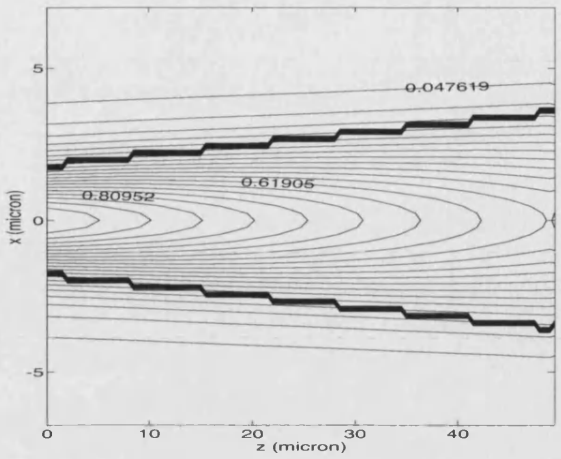
The comparison of the computational runtimes shows that the present method is faster than both the FD-BPM and LME methods. More specifically, in the case of FD-BPM and LME the structure is divided longitudinally in N sections (the larger the number N the more accurate being the results) and the field is computed for each section. In the case of the present method the field is computed overall the entire device in only one computer cycle. Therefore the FD-BPM and LME require a computational time of the order N times longer than the present method.

4.8.2 Effect of taper angle on propagation

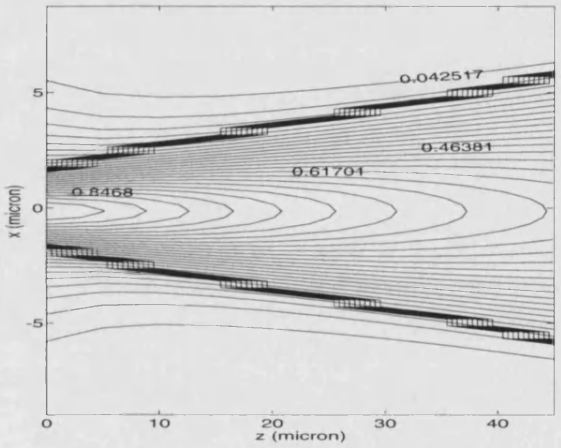
Optical propagation in different angle tapers are solved in order to examine how the same field launched at the start of the taper propagates in the structure. Three taper angles are chosen for comparison, $\theta_0 = 0.1^o$, 2^o, 5^o, each small enough to ensure that the paraxial approximation is valid.

When the angle is small, $\theta_0 = 0.1^o$, Fig. [4.7a], the field approaches the propagation in uniform media. Conversely, with a large angle, $\theta_0 = 5^o$, the effects of the taper are more evident: in one hand the waist of the field increases as the waist of the rib is increased, but, on the other hand, the taper enforces more lateral confinement in the bottom region, Fig. [4.7c].

The dependence of the field propagation on the taper angle can be quantified by considering the values obtained for the parameter q in the slowly varying term $a(z)$, eqn. (4.6). Tab. [4.2] describes how larger angles require higher correction terms to the chosen fast propagation term $p = \beta$. The values of q act on both amplitude and phase of the slowly varying term, $F(x, y, z)$. However, the effect on the phase is small compared to the fast variation, so that the slowly varying approximation is valid: in fact, examining the values of q_R in Tab. [4.2] it is always $q_R/p \ll p$. Regarding the amplitude of $a(z)$, eqn. (4.7), the decay is larger in taper with larger angle, i.e. amplitude of q_I increases with the angle. Since in a large angled taper the field waist broadens faster the amplitude decay must follow accordingly in order to ensure the conservation of transmitted energy at any transverse cross-section, ($z = \text{const}$).



(a)

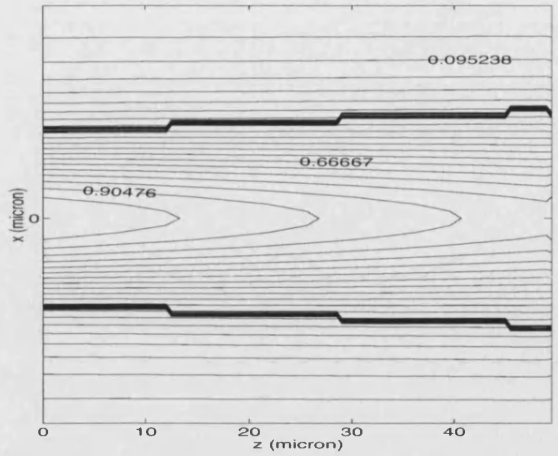


(b)

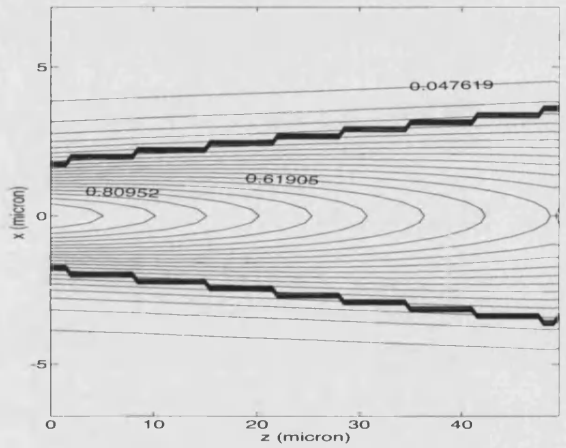
Figure 4.6: Comparison of results with LME method. Field profile contour plot, in Arbitrary Units (A.U.), at the longitudinal cross section underneath the rib, $y = -0.5\mu m$, Fig. [4.1]. (a) present model; (b) Local Mode Expansion method. Angle $\theta_0 = 2^o$.

$p = 14.5197$		
angle	q_R	q_I
0.5^o	0.257	0.054
2^o	0.298	0.149
5^o	0.308	0.221

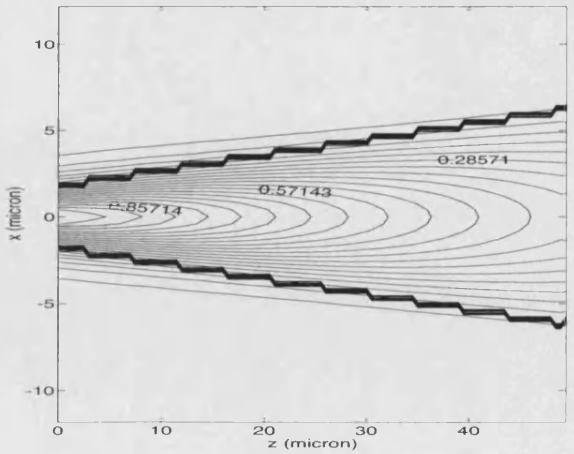
Table 4.2: Values obtained for $q = q_R + iq_I$ at different taper angles.



(a)



(b)



(c)

Figure 4.7: Field propagation for different taper angles. Field profile contour plot, in Arbitrary Units (A.U.), at the longitudinal cross-section $y = -0.5\mu m$. (a) $\theta_0 = 0.1^o$; (b) $\theta_0 = 2^o$; (c) $\theta_0 = 5^o$.

4.8.3 Total longitudinal phase variation

An indirect test for the validity of the model can be pursued by examining the total longitudinal phase variation along the z direction. According to the expression utilised for the field inside the rib region, eqns. (4.2), (4.5), (4.6), the longitudinal phase variation is given by the contribution of two terms: (a) the fast propagating term p, (b) the phase part of the slowly varying term $a(z)$, i.e. q_R/p. Therefore, the total longitudinal phase variation is

$$\phi_{tot} = p + q_R/p \tag{4.17}$$

In the limit case of uniform waveguide, Fig. [4.8a], (flare angle $\theta_0 = 0^o$) the longitudinal phase constant is equivalent to the propagation constant, $\phi_{tot} = \beta$. The opposite limit case, with flare angle $\theta_0 = 90^o$, Fig. [4.8b], is equivalent to the propagation of the initial field $E_0(x, y)$ into the half space region $z > 0$, made of the vertical layer sequence cladding-core-substrate. This sequence constitutes a longitudinally uniform slab waveguide and, hence the total phase ϕ_{tot} can not exceed the effective propagation constant of such slab waveguide. The total phase variation for all other angles lie between these two limit cases, Fig. [4.9].

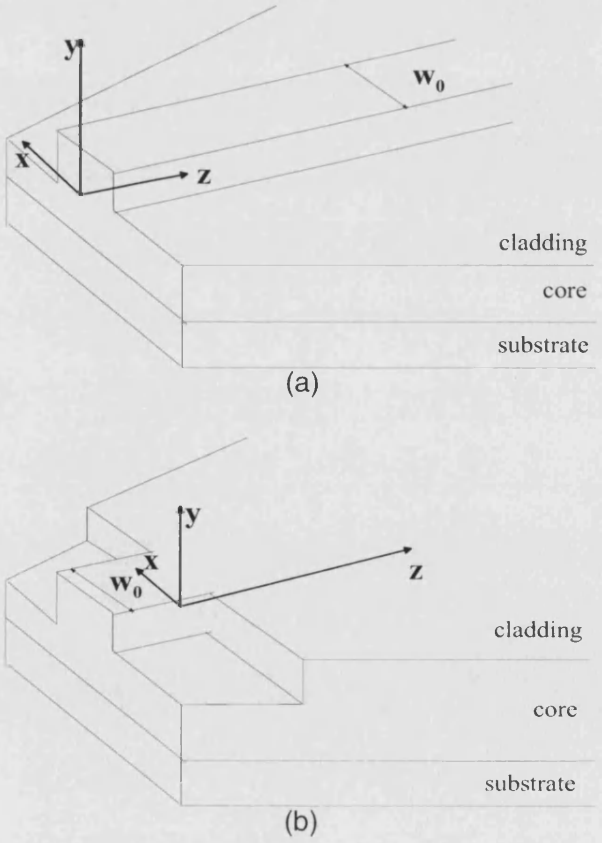


Figure 4.8: Limit cases for the flaring angle. (a) $\theta_0 = 0^o$; (b) $\theta_0 = 90^o$

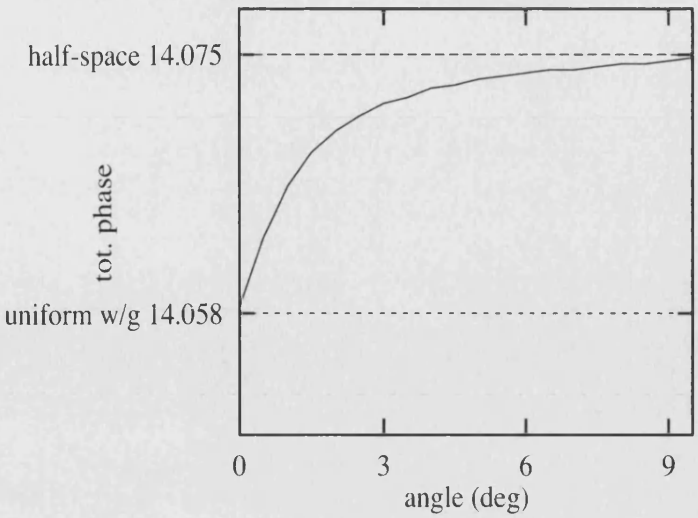


Figure 4.9: Total phase, ϕ_{tot}, of the propagating field for different flare angles. The total phase varies between the uniform waveguide propagation constant and the half-space phase propagation (asympotically).

4.9 Y-junction waveguide

The method presented in this chapter for solving tapered devices can be extended to Y-junction configurations, Fig. [4.10]. The inclusion of the two arms is similar to the extension of the Spectral Index Method for a longitudinally uniform single rib waveguide to a two ribs waveguide coupler, presented in Section 3.5.1.

In the present case each mode, symmetric and anti-symmetric, is launched separately

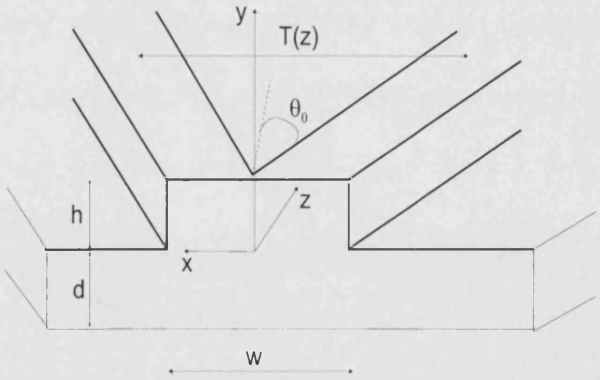


Figure 4.10: Structure of a Y-junction waveguide.

from the uniform input rib waveguide at the start of the Y-junction, $z = 0$. The total optical field propagating along the two arms is then obtained by adding the two field profiles, calculated separately. The field inside the top part is approximated with the expression

$$E_A(x, y, z) = F_A(x, y, z) \cdot \exp(-ipz) = a(z) \cdot f(x, z) \cdot G(y) \cdot \exp(-ipz) \qquad (4.18)$$

with

$$f(x,z) = \begin{cases} A_1 \cos\left[\frac{(2k-1)\pi}{w}(x+T/2)\right] & (-w-T)/2 < x < (w-T)/2 \\ A_1 \cos\left[\frac{(2k-1)\pi}{w}(x-T/2)\right] & (T-w)/2 < x < (w+T)/2 \\ 0 & elsewhere \end{cases} \quad \text{for symmetric fields}$$

$$f(x,z) = \begin{cases} A_1 \cos\left[\frac{(2k-1)\pi}{w}(x+T/2)\right] & (-w-T)/2 < x < (w-T)/2 \\ A_1 \sin\left[\frac{2k\pi}{w}(x-T/2)\right] & (T-w)/2 < x < (w+T)/2 \\ 0 & elsewhere \end{cases} \quad \text{for anti-symmetric fields} \tag{4.19}$$

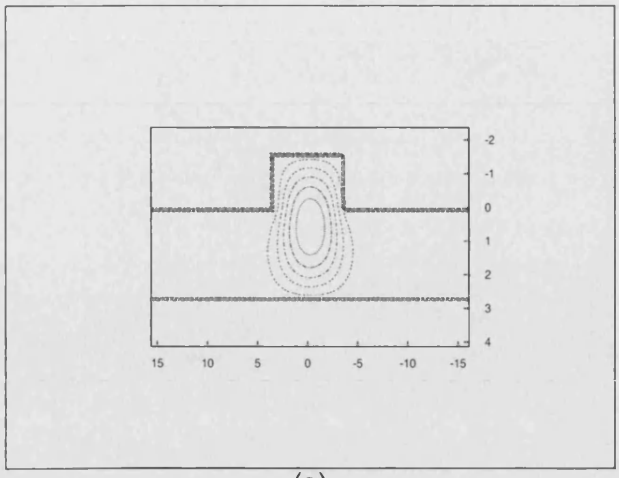
The slowing varying term $a(z)$ is determined in Appendix D, and given by

$$a(z) = \exp\left(-i\frac{q}{p}z\right) \qquad q = q_R - iq_I \tag{4.20}$$

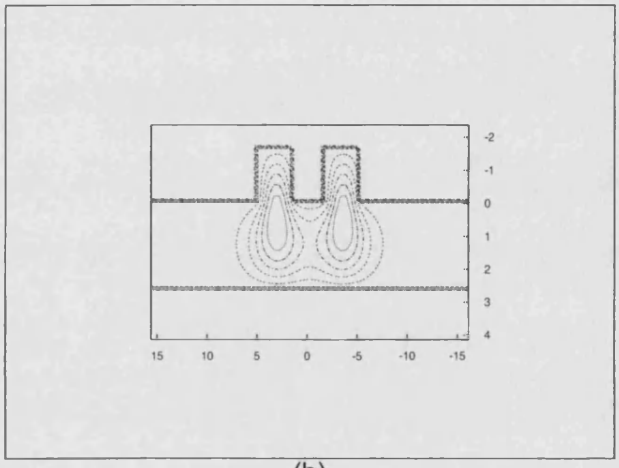
q_R describes the small phase correction to the arbitrarily chosen fast propagation p, and q_I the amplitude decrease as the field propagates longitudinally.

The field in the region below the rib is Fourier transformed along x, the same as described in eqns. (4.8), (4.9) for the tapered geometry. The boundary/interface conditions and the variational expression are similar to those applied in the taper waveguide, i.e. eqns. (4.14), (4.15), (4.16).

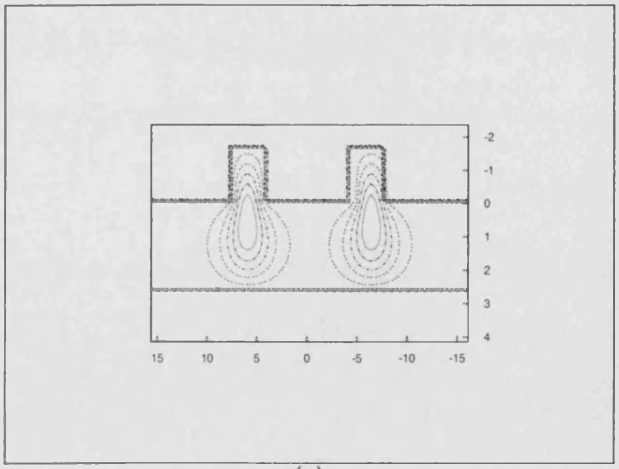
An example of field propagation, for a Y-Junction rib waveguide with parameters as given in Tab. [4.1] and an angle $\theta_0 = 2^o$, is shown in Figs. [4.11], [4.12]. A symmetric field is launched at the start of the Y-junction, from the uniform input waveguide, Fig. [4.11a]. The field splits and propagates equally between the two arms, decreasing in amplitude, Fig. [4.12], because of coupling to other modes (radiative). Although the radiated modes are not explicitly accounted for, their effect is included implicitly in the variational expression (4.16), so that the slowly varying term $a(z)$ is determined accordingly.



(a)

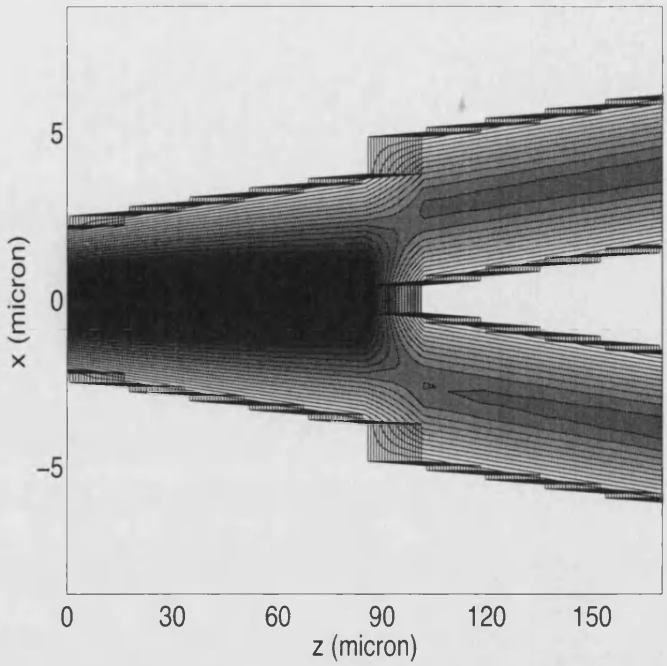


(b)

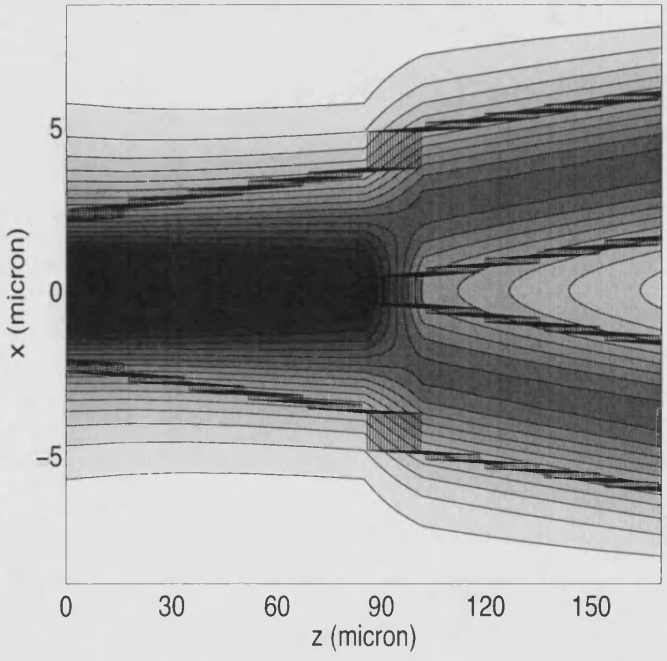


(c)

Figure 4.11: Y-junction structure. Field profile contour plot, in Arbitrary Units (A.U.), at different transverse cross sections, Fig. [4.10]. (a) $z = 0$, initial field fed by the input rub waveguide; (b) $z = 30\mu m$; (c) $z = 66\mu m$. Split angle, $\theta_0 = 2^o$.



(a)



(b)

Figure 4.12: Y-junction structure. Field profile contour plot, in Arbitrary Units (A.U.), at different longitudinal cross sections, Fig. [4.10]. (a) top rib region, $y = 0.5\mu m$; (b) region below the rib, $y = -0.5\mu m$. Split angle, $\theta_0 = 2^o$.

4.10 Quasi-modal propagation in taper rib waveguides

The Spectral Index Method determines, under some approximations, the eigenmode solutions of longitudinally uniform rib waveguides, by expressing the field inside the rib region in terms of metal-like rectangular waveguide eigenmodes. In fact, because of the effective rib with zero electric field outside, Fig. [3.2], the two vertical walls of the rib can be seen as parallel metal plates, which admits eigenmodes of the type described in eqn. (3.6).
In the case of a taper waveguide, Fig. [4.2], the effective rib is equivalent to a metal-like horn waveguide. Since metal horns admit eigenmode solutions it could be argued that, by applying the same concepts as in the SIM, the field inside a taper deep ridge could be expressed in terms of metal horn eigenmodes. However, if by using metal horn eigenmodes it is still possible to determine eigenmode type solutions in a taper rib waveguide is not entirely clear, and certainly needs more investigation.
In this section a preliminary analysis is attempted to investigate the possibility of solving taper waveguides in cylindrical coordinates, by applying metal horn eigenmodes. The theory presented here does not aim to be exhaustive, but mainly to introduce material for possible further developments and discussions.

4.10.1 Mode solutions of a metal horn

As introduced in Section 2.5, the distinct feature of eigenmode solutions is the fact that they can be expressed as products of factors with separated variables, i.e. the transverse, x and y axes are separated from the longitudinal z axis, eqn. (2.14). This feature is valid even for fields propagating in a metal horn of sectoral shape, Fig. [4.13]. In fact, the field in metal horns can be expressed as separate factors of the transverse axes, θ and y, and the radial direction, r, [6]. However, in order to obtain an analytic solution for the eigenmodes of a sectoral horn it is necessary to solve the general wave equations (2.6), (2.7) in cylindrical coordinates, (r, θ, y).

As explained in Section 2.4, the electromagnetic field in cylindrical coordinates can be easily solved, in scalar form, if eqn. (2.6) is solved for the vertical component of the electric field, E_y, or eqn. (2.7) is solved for the vertical component of the magnetic field, H_y, (TM_y and TE_y polarisation, respectively). For TM_y polarisation the wave equation is, thus,

$$\frac{\partial^2 E_y}{\partial r^2} + \frac{1}{r}\frac{\partial E_y}{\partial r} + \frac{1}{r^2}\frac{\partial^2 E_y}{\partial \theta^2} + \frac{\partial^2 E_y}{\partial y^2} + {k_0}^2 \epsilon E_y = 0 \tag{4.21}$$

Solutions of eqn. (4.21), with reference to Fig. [4.13], are, [7], [6]:

$$E_{mn}(r, \theta, y) = \cos\left[\frac{(2m+1)\pi}{\theta_0}\theta\right] \cdot \cos\left(\frac{n\pi}{d}y\right) \cdot \left\{ \begin{array}{c} H_\nu^{(1)}(k_n r) \\ H_\nu^{(2)}(k_n r) \end{array} \right\} \tag{4.22}$$

the two indices, $m = 0, 1, 2, \ldots$ and $n = 1, 2, \ldots$, are the azimuthal and the vertical mode numbers, respectively; $H_\nu^{(1)}$, $H_\nu^{(2)}$ are the reverse and forward travelling Bessel functions

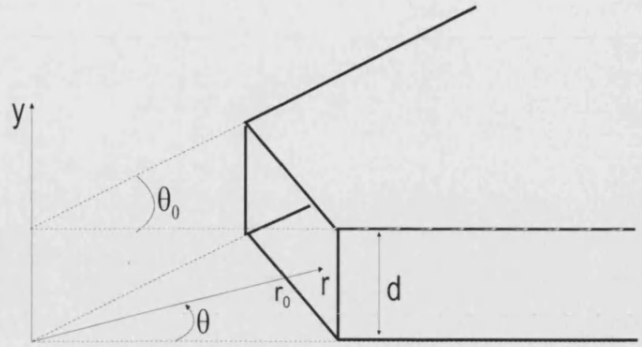


Figure 4.13: Structure of a sectoral metal horn.

of the third type (or Hankel functions); $\nu = \frac{(2m+1)\pi}{\theta 0}$ indicates that the Hankel functions are of fractional order; $k_n^2 = k_0^2\epsilon - (\frac{n\pi}{d})^2$ is the radial wavenumber. The set of eqns. (4.22) are orthogonal and complete and are not valid at the apex, $r = 0$.

4.10.2 Field expression inside the rib

Similarly to the SIM for uniform waveguides, the analysis of the field in taper rib waveguides is subdivided in two regions: inside the taper rib and below the rib.
After replacing the actual rib with an *effective* taper rib, thus introducing metal-like boundary conditions, the field, E_A, inside the rib can be expanded, along the azimuthal axis, in terms of metal horn modes, eqn. (4.22),

$$E_A = \sum_{m=1}^{\infty} \cos\left[\frac{(2m+1)\pi}{\theta_0}\theta\right] \cdot G(y) \cdot \left[a_1 H_\nu^{(1)}(kr) + a_2 H_\nu^{(2)}(kr)\right] \qquad \nu = \frac{(2m+1)\pi}{\theta_0} \quad (4.23)$$

In the case of a metal horn the forward and reverse travelling solutions are not coupled, so that they can be used separately and independently from each other to represent an exact mode solution of the horn. In the present case of taper rib waveguide, since the rib region does not have metal-like boundary conditions at the bottom surface, $y = 0$, Fig. [4.2], the forward and reverse travelling waves are, in general, coupled. However, because the analysis is focused only on adiabatic tapers, it is assumed that the coupling is small so that the forward and reverse travelling waves are still treated separately. Moreover, as first approximation, only one term of the expansion (4.23) is considered at a time. Thus, for $m = 1$, the electric field inside the rib region is

$$E_A = \cos\left(\frac{\pi}{\theta_0}\theta\right) \cdot G(y) \cdot H_\nu^{(2)}(kr) \qquad \nu = \frac{\pi}{\theta_0} \quad (4.24)$$

with k the radial wavenumber to be determined.
Inserting the expression (4.24) in the wave eqn. (4.21) gives

$$\frac{\mathrm{d}^2 H_\nu^{(2)}}{\mathrm{d}r^2} G(y) + \frac{1}{r}\frac{\mathrm{d}H_\nu^{(2)}}{\mathrm{d}r} G(y) + \left(k_0^2\epsilon - \frac{\nu}{r^2}\right) H_\nu^{(2)} G(y) + \frac{\mathrm{d}^2 G(y)}{\mathrm{d}y^2} H_\nu^{(2)} = 0 \tag{4.25}$$

utilising separation of variables conditions, yields

$$\frac{\mathrm{d}^2 H_\nu^{(2)}(kr)}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}H_\nu^{(2)}(kr)}{\mathrm{d}r} + \left(k_0^2\epsilon - \alpha^2\right) H_\nu^{(2)}(kr) - \frac{\nu}{r^2} H_\nu^{(2)}(kr) = 0 \tag{4.26}$$

$$\frac{\mathrm{d}^2 G(y)}{\mathrm{d}y^2} = -\alpha^2 G(y) \tag{4.27}$$

Eqn. (4.26) is the Bessel equation, with $k^2 = k_0^2\epsilon - \alpha^2$.
Eqn. (4.27), together with the boundary condition at the top of the rib, $E_A(r, \theta, y = h) = 0$, Fig. [4.2], provides the expression for $G(y)$ as

$$G(y) = \frac{\sin \alpha(h - y)}{\sin \alpha h} \tag{4.28}$$

4.10.3 Field expression in the region below the rib

The region below the rib, $y < 0$, is uniform along the radial and azimuthal axes. A general field expansion is:

$$E_B = \sum_{m=0}^{+\infty} A_m(r, y) \cos(m\theta) \tag{4.29}$$

Inserting eqn. (4.29) in the cylindrical wave equation (4.21), and considering the fact that the cosine functions are orthogonal, produces

$$\frac{\partial^2 A_m}{\partial r^2} + \frac{1}{r}\frac{\partial A_m}{\partial r} + \left(k_0^2\epsilon - \frac{m^2}{r^2}\right) A_m + \frac{\partial^2 A_m}{\partial y^2} = 0 \tag{4.30}$$

for $m = 0, 1, 2, \ldots$.
The functions $A_m(r, y)$ can be expressed by using the Hankel transform, [8],

$$A_m(r, y) = \int_{0+}^{+\infty} k_m G_m(k_m, y) \cdot H_\nu^{(2)}(k_m r) \mathrm{d}k_m \qquad \nu = \frac{\pi}{\theta_0} \tag{4.31}$$

Inserting the expansion (4.31) in eqn. (4.30), and interchanging the integral and the derivative operators, yields

$$\int_{0+}^{+\infty} k_m \left[\frac{\mathrm{d}^2 H_\nu^{(2)}}{\mathrm{d}r^2} G_m + \frac{1}{r}\frac{\mathrm{d}H_\nu^{(2)}}{\mathrm{d}r} G_m + \left(k_0^2\epsilon - \frac{m^2}{r^2}\right) G_m H_\nu^{(2)} + \frac{\partial^2 G_m}{\partial y^2} H_\nu^{(2)}\right] \mathrm{d}k_m = 0 \tag{4.32}$$

Because of Bessel's equation

$$\frac{\mathrm{d}^2 H_\nu^{(2)}(k_m r)}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}H_\nu^{(2)}(k_m r)}{\mathrm{d}r} = -\left(k_m^2 - \frac{\nu^2}{r^2}\right) H_\nu^{(2)}(k_m r) \tag{4.33}$$

eqn. (4.32) can be rewritten as

$$\int_{0+}^{+\infty} k_m \left[(k_0^2\epsilon - k_m^2) G_m H_\nu^{(2)} + \left(\frac{\nu^2 - m^2}{r^2}\right) G_m H_\nu^{(2)} + \frac{\partial^2 G_m}{\partial y^2} H_\nu^{(2)} \right] \mathrm{d}k_m = 0 \qquad (4.34)$$

Eqn. (4.34) can be simplified by applying the orthogonality condition, [9],

$$\int_a^b r \cdot H_\nu^{(2)}(k_m r) H_\nu^{(2)}(\bar{k}_m r) \mathrm{d}r = \begin{cases} 0 & k_m \neq \bar{k}_m \\ I(\bar{k}_m) & k_m = \bar{k}_m \end{cases} \qquad (4.35)$$

with $0 < a < b$ and

$$I(k_m) = \left[\frac{r^2}{2} \{ H_\nu'^{(2)}(k_m r) + (1 - \frac{\nu^2}{k_m r^2}) H_\nu^{(2)}(k_m r) \} \right]_a^b \qquad (4.36)$$

Therefore, multiplying eqn. (4.34) by $H_\nu^{(2)}(\bar{k}_m r)$, integrating between the start of the taper $r = r_0$ and $r = +\infty$ and using the orthogonality eqn. (4.35), eqn. (4.34) becomes

$$\frac{\partial^2 G_m(\bar{k}_m, y)}{\partial y^2} I(\bar{k}_m) + (k_0^2\epsilon - \bar{k}_m^2) I(\bar{k}_m) G_m(\bar{k}_m, y) + \frac{1}{\bar{k}_m} \int_{0+}^{+\infty} k_m(\nu^2 - m^2) J(k_m) G_m(k_m, y) \mathrm{d}k_m = 0 \qquad (4.37)$$

with

$$J(k_m) = \int_{r_0}^{\infty} \frac{1}{r} \cdot H_\nu^{(2)}(k_m r) H_\nu^{(2)}(\bar{k}_m r) \mathrm{d}r \qquad (4.38)$$

The integral in eqn. (4.37) represents the coupling between different terms, $G_m(k_m y)$, due to the presence of the top sectoral rib.

In the set of eqns. (4.37) the functions $G_m(\bar{k}_m, y)$ and the constants $\bar{k}_m$ must be determined, for $m = 0, 1, 2, \ldots$. In principle, such unknown quantities can be determined by solving the interface/boundary conditions and by assigning an initial field profile for the propagating field. However, it is cumbersome to determine a direct quasi-analytic link between the initial field profile and the values of G_m and $\bar{k}_m$.

In order to maintain a quasi-analytic approach to the problem an approximation for the field expression in the region below the rib is invoked. Each coefficient, $A_m(r, y)$, is approximated by using only one spectral term, $G_m(k_m, y)$, of the Hankel Transform (4.31). In particular, the term corresponding to the radial wavenumber k of the top region, eqn. (4.24), i.e. $k^2 = k_0^2\epsilon - \alpha^2$, is chosen. With this choice the coefficients $A_m(r, y)$ in eqn. (4.31) become

$$A_m(r, y) = G_m(k, y) H_\nu^{(2)}(kr) \qquad (4.39)$$

and, consequently, the field expansion (4.29) is

$$E_B = \sum_{m=0}^{+\infty} G_m(k, y) \cos(m\theta) H_\nu^{(2)}(kr) \qquad (4.40)$$

It is important to note that although in each term of the expansion (4.40) the transverse spatial variables are separated, the entire expression of the field, E_B, does not have separated variables. The expression for $G_m(k, y)$ is determined from eqn. (4.37), which, with

the present assumption, i.e. $k_m = k$, reduces to

$$\frac{\partial^2 G_m}{\partial y^2} I(k) + (k_0^2 \epsilon - k^2) I(k) G_m + (\nu^2 - m^2) J(k) G_m = 0 \tag{4.41}$$

Solutions of eqn. (4.41) are

$$G_m(k,y) = \begin{cases} B_1 \sin(k_{ym} y) + B_2 \cos(k_{ym} y) & -d < y < 0 \quad \text{core} \\ C_1 \exp(\Gamma_{ym} y) & y < -d \quad \text{substrate} \end{cases} \tag{4.42}$$

with k_{ym} and Γ_{ym} given by:

$$\begin{aligned} k_{ym}^2 &= k_0^2 \epsilon_2 - k^2 + (\nu^2 - m^2) \tfrac{J(k)}{I(k)} \\ \Gamma_{ym}^2 &= k^2 - k_0^2 \epsilon_3 - (\nu^2 - m^2) \tfrac{J(k)}{I(k)} \end{aligned} \tag{4.43}$$

and B_1, B_2, C_1 coefficients to be determined by applying the boundary conditions at the common interface $y = -d$. The solution in the substrate, $y < -d$, is such that the field is guided inside the core and vanishes at $y = -\infty$.

The continuity of the field is also ensured at the bottom of the rib, $y = 0$, i.e., by using eqn. (4.24), (4.28), (4.40),

$$E_A(r,\theta,y=0) = \cos(\frac{\pi}{\theta_0}\theta) H_\nu^{(2)}(kr) = \sum_{m=0}^{+\infty} G_m(k,y=0)\cos(m\theta) H_\nu^{(2)}(kr) = E_B(r,\theta,y=0) \tag{4.44}$$

Similarly to the SIM for uniform waveguides, the continuity of the field derivative can not be ensured (Section 3.4), and a variational analysis is therefore applied

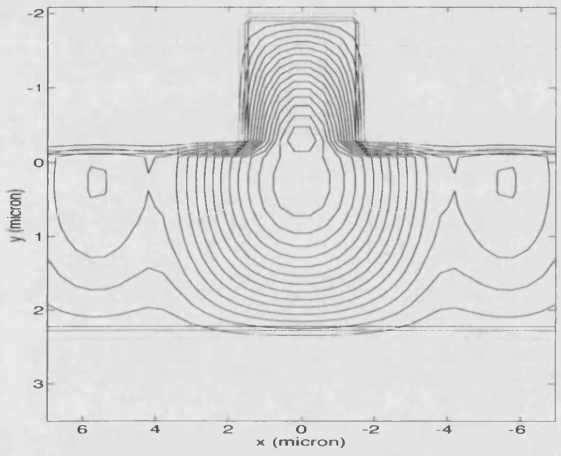
$$\int_{-\pi}^{+\pi}\int_{r_0}^{+\infty} \left[E_A^* \frac{\partial E_A(r,\theta,y=0^+)}{\partial y} - E_B^* \frac{\partial E_B(r,\theta,y=0^-)}{\partial y} \right] \mathrm{d}\theta \mathrm{d}r = 0 \tag{4.45}$$

The main difference between the variational expression (4.45) and the expression used in (4.16) to solve the taper waveguide propagation in Cartesian coordinates, is that the second term in (4.16) is not needed in the present case. In fact, the energy conservation is ensured automatically by the type of function chosen to describe the longitudinal propagation, i.e. Hankel function, eqn. (4.24), rather than plane wave function, eqn. (4.2). In Cartesian coordinates the variational analysis is applied to minimise the field derivative mismatch at the base of the rib and to ensure energy conservation as the field propagates longitudinally. In Cylindrical coordinates the variational analysis is invoked only to minimise the mismatch of the field derivative at the base of the rib. Thus, for taper waveguides, the choice of Hankel functions to represent the field inside the rib is more appropriate than to approximate the field with sine/cosine plane waves as in Cartesian coordinates. Moreover, in the Cartesian coordinates case the fast propagating term, p, was chosen to assign the initial propagating field, and the variational expression (4.16) was used to obtain the (complex) slowly varying term, $a(z)$. Differently, the variational expression (4.45) is solved to determine the value of the (real) radial propagation term, k.

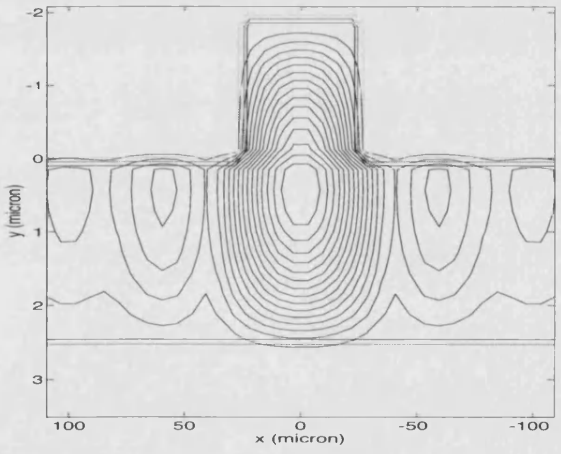
4.10.4 Results: optical field distribution

The taper waveguide with parameters as in Tab. [4.1] is solved in cylindrical coordinates. Some preliminary results of the field profile are shown in Fig. [4.14]. Similarly to the field profiles determined in Cartesian coordinates, Fig. [4.3], the effect of the taper produces an increasing effective lateral confinement as the field propagates.

The side lobes appearing in Figs. [4.14] are generated because the expansion (4.40)



(a)



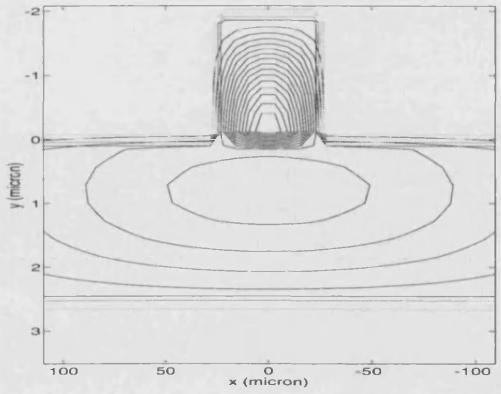
(b)

Figure 4.14: Field profile contour plot, in Arbitrary Units (A.U.), at different transverse cross sections, Fig. [4.1]: (a) $r = 40\mu m$; (b) $r = 400\mu m$. Angle $\theta_0 = 2.5^o$.

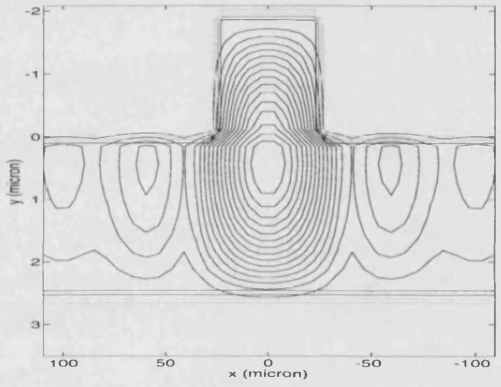
has been truncated and the field represented with a finite number of terms, $N = 50$. Fig. [4.15] shows how an expansion with too few terms generates a profile which does not satisfy accurately the continuity of the field at the base of the rib, $y = 0$, with consequent

excessive lateral spreading of the field in the bottom region.

If, in the limit case, the depth, d, of the uniform core region, Fig. [4.2], is reduced to a small



(a)



(b)

Figure 4.15: Influence of the number of terms, N, in the field expansion. Field profile contour plot, in Arbitrary Units (A.U.), at the transverse cross sections, $r = 50\mu m$: (a) N=10; (b) N=50. Angle $\theta_0 = 2.5^o$.

value, $d \to 0$, and the core/substrate refractive index step is chosen large, i.e. $(\eta_2 - \eta_3)$ is large, then the taper rib structure resembles a pure metal horn. Results for such a case are depicted in Fig. [4.16]. The field profile calculated is similar to a metal horn eigenmode solution, eqn. (4.22), with the peak centred in the middle of the rib.

The above results on the field propagation solved in cylindrical coordinates are presented at a preliminary stage and need to be analysed more. In particular, it is necessary to determine the formal connection between such results and those obtained by solving the same taper structure in Cartesian coordinates. Nevertheless, the first inspection reveals

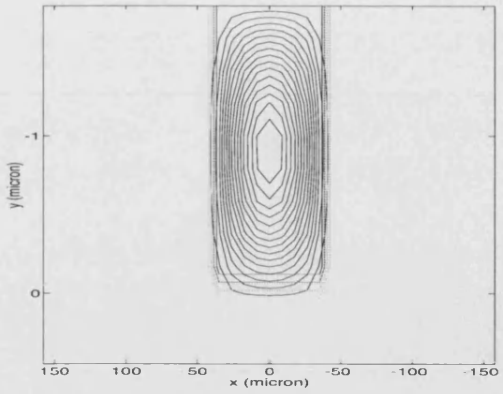


Figure 4.16: Limit case structure: taper rib as metal horn. Field profile contour plot, in Arbitrary Units (A.U.), at the transverse cross sections, $r = 50\mu m$.

qualitative satisfactory agreement between the profiles obtained by the two methods.

Bibliography

[1] März R., "Integrated Optics. Design and Modeling", Artech House, London, 1995

[2] Sewell P., Benson T.M., Greedy S., Vukovic A., "Advances in the Spectral Index Method for OEIC Design", *International conf. On Transparent Optical Networks*, 9-11 June, Kielce, Poland, 1999, 137-138

[3] Robson P.N., Kendall P.C., "Rib Waveguide Theory by the Spectral Index Method", RSP, England, 1990

[4] Marcuse D., "Light Transmission Optics", Van Nostrand Reinhold, 1982

[5] Hewitt P.D., Reed G.T., "Improving the Response of Optical Phase Modulators in SOI by Computer Simulation", *J. Lightwave Technol.*, V.18, No.3, 2000, 443-450

[6] Barrow W.L., Chu L.J., "Theory of the Electromagnetic Horn", *Proceeding of IRE*, 1939, 51-64

[7] Harrington R.F., "Time-Harmonic Electromagnetic Fields", McGraw-Hill, New York, 1961

[8] Tranter C.J., "Bessel Functions with Some Physical Applications", The English Universities Press, London, 1968

[9] Abramovitz M., Stegun I., "Handbook of Mathematical Functions", Dover Publications Inc., New York, 1965

Part II

Injection of carriers in semiconductor devices

Chapter 5

Injection of carriers into semiconductors

The first part of the thesis was devoted to describe optical propagation along planar semiconductor rib devices. An important requirement for optoelectronic components is the capability of altering their optical propagation characteristics. One relevant example is given by the optical switch, whose functional purpose is to select the spatial path of the optical signal. Similarly, a spatial modulator, for phase/amplitude modulation, requires alteration of the optical propagation characteristics.
This requirement can be achieved by changing the refractive index of the material thus affecting the optical properties of the device. The refractive index of a material can be altered, for example, during the growing process, adding impurities inside the semiconductor lattice to create a new alloy composition. Although this is the most effective way to influence the refractive index, it is not convenient because it can be only achieved during material growth.
However, alteration of the refractive index while the device is operating is more useful. One possible way of changing the refractive index of a semiconductor device is by injecting carriers (holes and electrons). It is then useful to determine the carrier distribution (which in general may not be uniform) inside the device after injecting an externally applied current.

5.1 Refractive index change in semiconductors

There are three most common different ways of altering the refractive index, η, of a semiconductor material:

- applying an external electric field (Electrooptic effect)
- injecting/depleting a current (Carriers induced effect)
- altering the temperature of the semiconductor (Thermooptical effect)

The effectiviness of these methods varies accordingly to the material utilised.

Electrooptic effect (or electrorefraction) is determined by two physical effects, both generated by an externally applied electric field: the Franz-Keldysh and the Kerr effects. The inherent on/off response time is very small (generally sub-picosecond), and limited only by the RC time constant of the device. The modulation depends of the electric field polarisation.
In the particular case of Silicon material studies on electrorefraction have shown that the electrooptic effect is small, especially if compared with the other two methods, [1].

The **carrier induced effect** can be achieved either by injection, (negative $\Delta\eta$), or by depletion, (positive $\Delta\eta$), of carriers into the semiconductor. The carrier effect is determined by three main physical phenomena: 1) free carrier absorption; 2) bandfilling; 3) Coulombic interaction of carriers with impurities, [1], [2], [3], [4]. However, the major interest here is the combination of all of them, in order to describe the overall effects. For many applications it is an important advantage to alter the refractive index by carrier injection because it is independent of the polarisation of the optical signal.
In the injection case the response time is limited by the minority carrier lifetime ($\approx ns$ to μs), the depletion mode is expected to be faster because of carrier sweep out.
In Silicon the carrier injection effect has been empirically determined as, [1]:

$$\begin{aligned} \Delta\eta &\approx (\Delta N)^{0.8} \qquad \text{for holes} \\ \Delta\eta &\approx (\Delta N)^{1.05} \quad \text{for electrons} \end{aligned} \tag{5.1}$$

where ΔN is the injected carrier density.

Thermooptical modulation is in general the slowest way of altering the refractive index. In Silicon the depth of modulation is as effective as the carrier injection, i.e. $\Delta\eta = 2 \cdot 10^{-4}/K$, [5]. However, because of high thermal conductivity of Silicon it requires high power consumption, so that carrier injection still remains the preferred way of altering the refractive index.

A comparison of the refractive index change due to the three different methods is summarised in Tab. [5.1], at $\lambda = 1.3\mu m$ for Silicon material.

Carrier injection	
injection of holes, ΔN	
$\Delta N = 10^{18} cm^{-3}$	$\Delta\eta = -10^{-3}$
$\Delta N = 10^{19} cm^{-3}$	$\Delta\eta = -8 \cdot 10^{-3}$
$\Delta N = 10^{20} cm^{-3}$	$\Delta\eta = -5 \cdot 10^{-2}$
injection of electrons, ΔN	
$\Delta N = 10^{18} cm^{-3}$	$\Delta\eta = -6 \cdot 10^{-4}$
$\Delta N = 10^{19} cm^{-3}$	$\Delta\eta = -5 \cdot 10^{-3}$
$\Delta N = 10^{20} cm^{-3}$	$\Delta\eta = -7 \cdot 10^{-2}$
Electrorefraction	
Franz-Keldysh effect	
$E = 75kV/cm$	$\Delta\eta = 5 \cdot 10^{-6}$
$E = 100kV/cm$	$\Delta\eta = 1.3 \cdot 10^{-5}$
Kerr effect	
$\lambda = 1.3\mu m$	
$E = 100kV/cm$	$\Delta\eta = 10^{-6}$
$E = 1000kV/cm$	$\Delta\eta = 10^{-4}$
Thermooptical effect	$\Delta\eta = 2 \cdot 10^{-4}/K$

Table 5.1: Comparison of refractive index change, $\Delta\eta$, in Silicon by using electrorefraction, carrier injection and thermooptical effect, at $\lambda = 1.3\mu m$.

5.2 Injection of carriers into semiconductors

The process of injecting carriers into a semiconductor sample can be entirely described by the current flow continuity equations, together with Poisson' s equation, [6]. The injection of a current results in a net increase of carriers (hole and electron) in the volume of the sample. However, increase or decrease of carriers from the volume is not only due to the flow in-flow out, but also because some hole-electron pairs are generated while some recombine. If recombination (at the rate R for unit volume) or generation (at the rate G for unit volume) processes occur inside the region the net amount of generation per unit time per unit volume is $(G - R)$.

By enforcing hole and electron particle conservation separately yields the equations for the hole concentration distribution, $p(x, y, z, t)$, and the electron concentration distribution, $n(x, y, z, t)$, [8],

$$\frac{\partial p}{\partial t} = -\frac{1}{e}\nabla \cdot \mathbf{J_p} + (G - R)_p \tag{5.2}$$

$$\frac{\partial n}{\partial t} = \frac{1}{e}\nabla \cdot \mathbf{J_n} + (G - R)_n \tag{5.3}$$

with e the elementary charge constant. $\mathbf{J_p}$ and $\mathbf{J_n}$ are the hole and electron current densities, respectively. It is useful to note that p and n are scalar variables while the current densities $\mathbf{J_p}$ and $\mathbf{J_n}$ are vectors. The total injected current density is given by $\mathbf{J_{tot}} = \mathbf{J_p} + \mathbf{J_n}$.

If p_0, n_0, are the equilibrium concentrations, before injection, the injection of current introduces excess carriers concentrations, Δp, Δn, so that the final carrier concentrations, after injection, are

$$p = p_0 + \Delta p \qquad n = n_0 + \Delta n \tag{5.4}$$

The transport of carriers is due to diffusion and conduction phenomena, [7]:

$$\mathbf{J_p} = \mathbf{J_{pdiff}} + \mathbf{J_{pcond}} = e(-D_p \nabla p + p\mu_p \mathbf{E}) \tag{5.5}$$

$$\mathbf{J_n} = \mathbf{J_{ndiff}} + \mathbf{J_{ncond}} = e(D_n \nabla n + n\mu_n \mathbf{E}) \tag{5.6}$$

where D_p, D_n are the diffusion coefficients, μ_p, μ_n the carrier mobilities and $\mathbf{E}$ is the electric field applied across the semiconductor.

If the diffusion coefficients and the mobilities are constant substitution of the current eqns. (5.5), (5.6) in the continuity eqns. (5.2), (5.3) yields

$$\frac{\partial p}{\partial t} = D_p \nabla^2 p - \mu_p \nabla \cdot (p\mathbf{E}) + (G - R)_p \tag{5.7}$$

$$\frac{\partial n}{\partial t} = D_n \nabla^2 n + \mu_n \nabla \cdot (n\mathbf{E}) + (G - R)_n \tag{5.8}$$

The diffusion coefficients can be related to the mobilities through Einstein's relations, [6],

$$D_p = \frac{\mu_p kT}{e} \qquad D_n = \frac{\mu_n kT}{e} \tag{5.9}$$

with k Boltzmann's constant [1] and T the absolute temperature.
Finally, the electric field, $\mathbf{E}$, is related to the carrier densities through Poisson's equation

$$\nabla \cdot \mathbf{E} = \frac{e}{\epsilon}(p - n - p_0 + n_0) \tag{5.10}$$

In the case of an intrinsic region the concentrations at equilibrium are $p_0 = n_0 = n_i$; for p and n doped region it is $p_0 = N_a$ (acceptor concentration) and $n_0 = N_d$ (donor concentration), respectively.
From eqn. (5.10) it is also possible to observe that any electric field variation is present only in those spatial regions with a non zero net charge.

5.2.1 Recombination

Recombination is an important physical process in semiconductor materials contributing negatively to the net injection of carriers as it generates losses of carriers from the conduction and/or valence band. The number of carrier involved in this process, per unit volume per unit time, is described by the recombination rate $R(n, p, T)$ as a function of the carrier concentrations and the temperature.
In order to model all the different types of physical recombination it is possible to regard the recombination rate $R(n, p, T)$, for a fixed value of the temperature, as an expansion in power series, [6],

$$R(n, p, T) = c_1(T) + c_2(T)n + c_3(T)p + c_4(T)pn + c_5(T)p^2n + c_6(T)pn^2 + c_7(T)p^2n^2 + \dots \tag{5.11}$$

Each term of this expansion represents a different physical effect; terms of higher order represent higher order effects.
It is evident that the coefficient c_1 is zero, $c_1 \equiv 0$, because there can not be any recombination without the presence of any type of carrier. However, in order to represent recombination processes in eqns. (5.7), (5.8), it is convenient to represent all the different recombination terms in (5.11) in the same way. To each type of carrier is then associated a lifetime constant, τ, describing the time necessary to the carrier before recombining. Following this definition recombination rates can also be expressed as:

$$R_p = \frac{p}{\tau_p} \qquad R_n = \frac{n}{\tau_n} \tag{5.12}$$

The lifetime constants τ_p and τ_n effectively represent the overall recombination processes, and each term of eqn. (5.11) contributes, with different weights, to the values of τ_p and τ_n.

As a broad convenient classification recombination processes may be subdivided in two main categories: band-to-band and no-band-to-band recombination.

[1] $k = 8.617 \cdot 10^{-5} eV/K$

Band-to-band recombination

Band to band recombination occurs when an electron in the conduction band replaces the location of a hole in the valence band, at a lower energy level, relaxing an amount of energy equivalent to the difference between the two energy levels. It can be of two types: radiative and non-radiative.
In the radiative process the relaxed energy is radiated out. Because the process involves two carriers (electron and hole) it is represented through the term $c_4 pn$ in eqn. (5.11); c_4 is also known as bimolecular recombination coefficient B_r. Considering the representation introduced in eqn. (5.12)

$$R_p = B_r pn = \frac{p}{\tau_p} \qquad R_n = B_r pn = \frac{n}{\tau_n} \tag{5.13}$$

it is possible to estimate the effective radiative lifetime constants as:

$$\tau_p = \frac{1}{B_r n} \qquad \tau_n = \frac{1}{B_r p} \tag{5.14}$$

In the non-radiative case, also known as Auger recombination, the energy relaxed from the electron-hole pair is passed to a third carrier. The physical process involves three carriers, and it is thus represented by the two coefficients c_5, c_6, (known as Auger coefficients) in eqn. (5.11). By using the form (5.12) it is

$$R_p = C_p p^2 n = \frac{p}{\tau_p} \qquad R_n = C_n p n^2 = \frac{n}{\tau_n} \tag{5.15}$$

Consequently, the Auger lifetime is expressed in terms of the Auger coefficients C_p, C_n as:

$$\tau_p = \frac{1}{C_p pn} \qquad \tau_n = \frac{1}{C_n pn} \tag{5.16}$$

It is important to note that because R_p, R_n in eqn. (5.15) depend of the third power of the carrier concentration, Auger recombination becomes relevant only in the case of very large carrier injection, [9].

No band-to-band recombination

Carriers can also be lost because they are trapped by energy levels located inside the forbidden energy bandgap. These extra energy levels are often introduced by donors and acceptors dopant or by dislocation due to lattice imperfections. These types of recombination have been quantified and modelled by Shockley, Read and Moll as, [9],

$$R_{SRM} = \frac{np - {n_i}^2}{\tau_{p0}(n + n_i) + \tau_{n0}(p + n_i)} \tag{5.17}$$

The time constants τ_{p0} and τ_{n0} vary with doping densities and material, n_i^2 is the intrinsic (undoped) carrier concentration. Because only a single carrier is involved when trapped,

these recombination are represented by the coefficients c_2, c_3 in eqn. (5.11). Moreover, expressing the R_{SRM} coefficient in terms of lifetime constants, eqn. (5.12), it is

$$R_{SRM} = \frac{np - {n_i}^2}{\tau_{p0}(n + n_i) + \tau_{n0}(p + n_i)} = \frac{p}{\tau_p} + \frac{n}{\tau_n} \tag{5.18}$$

Another important type of recombination occurs near surfaces of semiconductor samples. In fact, surfaces introduce high densities of extra energy levels acting as recombination centres, [10]. Because one electron recombines with one hole, the flux of holes to the surface precisely equals the flux of electrons: thus there is not any net current flowing. Differently from the other types of recombinations (bulk recombination) surface recombination can not be expressed through the effective lifetime constant introduced in (5.12). They are quantified by a velocity coefficient, S_r, describing how fast carriers recombine at the surface.

According to the semiconductor material, some of the recombination processes introduced above are less relevant thus they may be neglected when modelling devices. Tab. [5.2] summarises the effects of the various recombination processes in the case of Silicon material. For each type of recombination it is reported the corresponding effective lifetime constant: according to eqn. (5.12) a shorter lifetime constant yields a larger amount of carriers recombining, R.
It is possible to observe that the most important recombination effect in Silicon is due to carriers trapping, while, in comparison, the radiative electron-hole recombination is almost negligible, even in the case of large amount of injected carriers. The Auger effect becomes relevant and must be included in the model only in the case of large carrier injection.

5.2.2 Levels of injection

The amount of carrier injection, or injection level, establishes which physical processes are dominant when injecting carriers and, therefore, must be considered when modelling carrier injection. A scheme of different injection levels for a typical n doped semiconductor is described in Fig. [5.1]. The case for a p type semiconductor is similar. Two main levels are relevant: the *low-level* injection and the *high-level* injection, [10].
At equilibrium, the majority carrier (electron) concentration, n_0, approximately equals the doping density, $n_0 \approx N_d$.
In the case of *low-level* injection the amount of injected carriers is much smaller than the doping density, i.e. $\Delta p \approx \Delta n \ll N_d$. Therefore, the minority carriers, $p = p_0 + \Delta p$, will be significantly changed, while the majority carriers, $n = n_0 + \Delta n$, although injected in the same amount as the minority, will remain almost unaffected, Fig. [5.1b], thus the resulting

Recombination coefficients of Si	
Radiative	$B_r = 10^{-15} cm^3/s$
Auger	$C_p = 1.8 \cdot 10^{-31} cm^6/s$, $C_n = 8.3 \cdot 10^{-32} cm^6/s$
R_{SRM}	$\tau_{p0} = 200ns$, $\tau_{n0} = 200ns$
Recombination lifetime	
Band-to-band (Radiative)	
$\Delta p = \Delta n = 10^{17} cm^{-3}$	$\tau_p = \tau_n \approx 10^{-2} s$
$\Delta p = \Delta n = 10^{19} cm^{-3}$	$\tau_p = \tau_n \approx 10^{-4} s$
Band-to-band (Auger)	
$\Delta p = \Delta n = 10^{17} cm^{-3}$	$\tau_p = \tau_n \approx 4 \cdot 10^{-4} s$
$\Delta p = \Delta n = 10^{19} cm^{-3}$	$\tau_p = \tau_n \approx 4 \cdot 10^{-8} s$
Trapping SRM	
any Δp, Δn	$\tau_{SRM} = \tau_{p0} + \tau_{n0} \approx 4 \cdot 10^{-7} s$

Table 5.2: Recombination coefficients for Silicon material and comparison between different recombination processes, [3], [9].

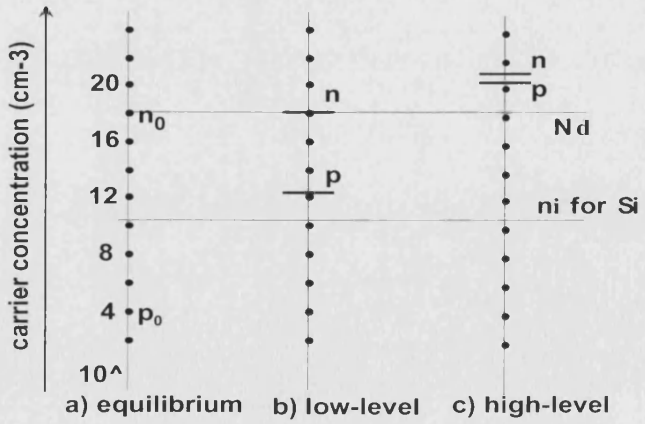


Figure 5.1: Levels of injection for a n-type Silicon material.

carrier concentration will be

$$n \approx N_d \qquad p \approx \Delta p \tag{5.19}$$

In the case of *high-level* injection, the excess injected carriers are comparable with the doping density. Therefore, both majority and minority carriers will be significantly affected, Fig. [5.1c].
Summarising, the conditions for the two cases of injection are:

- *low-level* injection: $\Delta p \approx \Delta n \ll$ majority carrier concentration

- *high-level* injection: $\Delta p \approx \Delta n$ almost equal or greater than majority carrier concentration

If the amount of injected carriers is small it is sufficient to solve the carrier eqns. (5.7), (5.8) only for the minority carriers. On the other hand, when the concentration of injected carriers is large it is necessary to solve for both the two types of carriers. This is usually the case when carriers are injected with the intention of changing the refractive index of the material. Indeed, in order to produce a considerable change it is necessary to inject a large amount of carriers, as shown in Tab. [5.1].

Bibliography

[1] Soref R.A., Bennett B.R., "Electrooptical Effects in Silicon", *IEEE J. Quantum Electron.*, QE-23, No.1, 1987, 123-129

[2] Tamir T., "Guided-Wave Optoelectronics", Springer-Verlag, Berlin, 1990

[3] Soref R.A., Lorenzo J.P., "All-Silicon Active and Passive Guided-Wave Components for $\lambda = 1.3$ and $1.6\mu m$", *IEEE J. Quantum Electron.*, V. QE-22, No.6, 1986, 873-879

[4] Bennett B.R., Soref R.A., Del Alamo J.A., "Carrier-Induced Change in Refractive Index of InP, GaAs, and InGaAsP", *IEEE J. Quantum Electron.*, V.26, No.1, 1990, 113-122

[5] Fisher U., Zinke T., Petermann K., "Integrated Optical Waveguide Switches in SOI", *Proceed. of the IEEE Intern. SOI Conference*, 2-5 Oct. 1995, Tucson, Arizona, S 141-142

[6] Adler R.B., Smith A.C., Longini R.L., "Introduction to Semiconductor Physics", John Wiley & Sons, NY, 1964

[7] Lampert M.A., Mark P., "Current Injection in Solids", Academic Press, NY, 1970

[8] McKelvey J.P., "Solid State and Semiconductor Physics", Harper & Row, NY, 1966

[9] Sze S.M., Physics of Semiconductor Devices, John Wiley & Sons, NY, 1981

[10] Grove A.S., "Physic and Technology of Semiconductor Devices", Wiley, 1967

Chapter 6

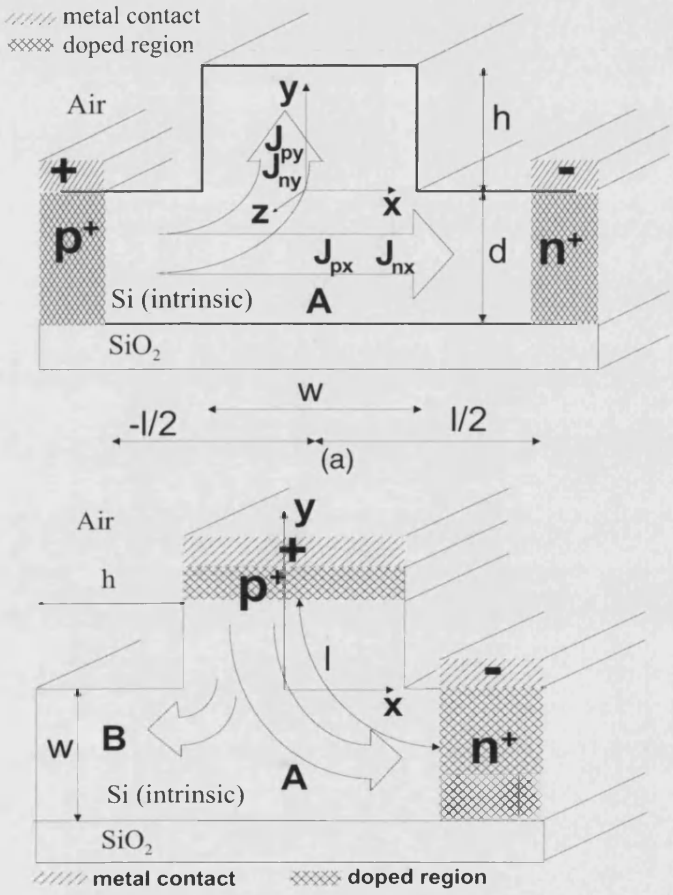
Injection of carriers in a rib waveguide

The general concepts of semiconductor carrier injection introduced in Chapter 5 are now applied specifically to a planar rib waveguide structure. The method presented for determining the two-dimensional carrier profile has been made simple to use and transparent to follow the 'physics in the computation', albeit at the expense of a slight loss of rigour. It is meant to provide a means for rapid initial evaluation of desirable design directions as is often required, for example, in an industrial context. The details of the method presented here applies to SOI stripe devices, but the method is applicable in a more general context upon minor adjustments appropriate to the case of interest.

6.1 Structure of the SOI device analysed

In Silicon-on-Insulator (SOI) devices the current can be injected only from the top surface, as the SiO_2 deposited in the bottom part of the Silicon optical guiding region is insulating. As a consequence, the two metal contacts for injecting current must necessarily be deposited on the top Si layer. Although this limitation, different configurations are still possible. If a rib is etched on the top layer, for example, the two metal contacts can be either deposited on the side of the rib or one contact on top and the other on the side of the rib, Fig. [6.1]. In the first case, Fig. [6.1a], the current flows mainly between the two side contacts, although some carriers diffuse inside the rib region. In the second case, Fig. [6.1b] there will be a distribution of carriers between the top and the side contacts mainly in the central region A; however, it is expected that a fraction of carriers will diffuse also into region B.

For the purpose of modelling carrier distributions the second structure can be treated similarly to the first; with the two metal contacts aligned in a straight line, on the side of a stripe of length w, separated by a distance l. Therefore, the amount of carriers that in



(b)

Figure 6.1: Different ways of depositing metal contacts in a SOI rib waveguide.

Fig. [6.1b] diffuses inside region B is represented by the amount of carriers that flows into the stripe in the equivalent structure of Fig. [6.1a]. Although this equivalence might seem too approximate it has the advantage of being straightforward.
However, in practice the configuration in Fig. [6.1b] is not very convenient for realising devices. In fact, in order to increase and facilitate the flux of current from a metal contact to the semiconductor region it is necessary to dope the region in proximity of the contact. The introduction of donor/acceptor generates an increase of optical losses. As shown in part one, the optical field is mainly confined in the central region below the rib, while only small tails extend on the two sides of the rib. Therefore, because of the optical confinement it is then more convenient to utilise the structure shown in Fig. [6.1a], in order to reduce the overlap between the optical field and the doped region, with consequent reduction of optical losses.
For this reason the carrier distributions analysis will henceforth be focused only to structures similar to that represented in Fig. [6.1a].

6.2 One-dimensional analysis: the p-i-n junction

As a first approximation the structure of Fig. [6.1a] can be regarded as a one-dimensional p-i-n structure, Fig. [6.2], ignoring the rib and assuming that the carrier distributions and the electric field are uniform along the y direction. Thus, in the steady state regime eqns. (5.7)-(5.10) can be rewritten, with the use of eqns. (5.5), (5.6), as

$$-D_p\frac{\mathrm{d}^2p(x)}{\mathrm{d}x^2}+\mu_p\frac{\mathrm{d}p(x)}{\mathrm{d}x}E(x)+\mu_p\frac{\mathrm{d}E(x)}{\mathrm{d}x}p(x)=(G-R)_p \tag{6.1}$$

$$-D_n\frac{\mathrm{d}^2n(x)}{\mathrm{d}x^2}-\mu_n\frac{\mathrm{d}n(x)}{\mathrm{d}x}E(x)-\mu_n\frac{\mathrm{d}E(x)}{\mathrm{d}x}n(x)=(G-R)_n \tag{6.2}$$

$$\frac{\mathrm{d}E(x)}{\mathrm{d}x}=\frac{e}{\epsilon}(p-n-p_0+n_0) \tag{6.3}$$

Because of their complexity this set of equations can not be solved analytically and hence a numerical approach is invoked. Many examples exist in published literature on how to determine carrier distributions in semiconductor devices. The seminal paper on the computation of a p-i-n junction was by Gummel, [1]. He suggested a solution based on an iterative method. First an educated guess for the electric field distribution is made: this assumption allows to decouple the two carrier eqns. (6.1), (6.2). Therefore the solution for the carrier concentrations can be obtained separately for $p(x)$ and $n(x)$. Finally the results of p and n are inserted back into Poisson's eqn. (6.3) which is then solved to determine the electric field distribution. The estimated profile $E(x)$, different from the initial guessed profile, is used to recalculate, with more accuracy, new distributions of $p(x)$ and $n(x)$. The

process is repeated iteratively, until the desired degree of accuracy is reached.

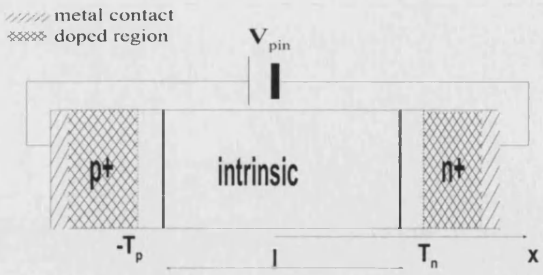


Figure 6.2: Structure of a p-i-n junction.

6.2.1 Boundary Conditions

In the present context the contacts of the p-i-n structure are located at infinite distance from the p-i and i-n junctions. The exact choice for the boundary conditions consists of assuming the electric field to be zero at the metal contacts, $E(-\infty) = E(+\infty) = 0$. Moreover, both the carrier concentrations, $p(x)$, $n(x)$, must reach their equilibrium value away from the junctions, i.e.,

$$p(-\infty) = p_0 = N_a \qquad n(+\infty) = n_0 = N_d \tag{6.4}$$

However, these conditions are applicable only if the problem is solved entirely analytically and the solutions are considered over an infinite spatial range. Conversely, when eqns. (6.1)-(6.3) are solved numerically these conditions can not be applied, since the infinite distance is replaced by an approximate large, but finite, distance (computational window). The accuracy of the solution will then depend on how far the metal contacts are chosen from the junctions. As a consequence the conditions (6.4) must be abandoned, and alternative, accurate yet realistic, boundary conditions need to be considered.
The approach followed in this work is based on the assumption that all the injected carriers recombine inside the intrinsic region before they reach the opposite junction [2]. Consequently, if holes are injected from the p side and electrons from the n side into the intrinsic region, Fig. [6.2], it follows that there will not be any hole current flowing through the i-n junction and, similarly, any electron current flowing through the p-i junction, [2],

$$\begin{aligned} J_n(x = -l/2) &= 0 \\ J_p(x = l/2) &= 0 \end{aligned} \tag{6.5}$$

Considering both the diffusive and the conductive currents, i.e. using eqns. (5.5), (5.6), the boundary conditions are expressed as

$$\begin{aligned} J_n(x = -l/2) &= e\mu_n n(-l/2)E(-l/2) + eD_n \tfrac{dn}{dx} = 0 \\ J_p(x = l/2) &= e\mu_p p(l/2)E(l/2) - eD_p \tfrac{dp}{dx} = 0 \end{aligned} \tag{6.6}$$

With the conditions (6.5) the boundaries are shifted from infinity to finite points of the device, corresponding to the junctions, although this choice limits the solution for $p(x)$ and $n(x)$ only within the intrinsic region, ($|x| < l/2$), Fig. [6.2].
However, since the investigation is focused mainly on the central region of the waveguide, where the optical field is mostly confined, it is sufficient, for the purpose of the present work to limit the analysis of carrier distributions only across the intrinsic region. For completeness, a quasi-analytic method for solving the carrier concentrations in the highly doped regions too is suggested in Appendix E. Although the doped regions may not be relevant, it is useful to solve for the carrier and electric field distributions also in these regions, in order to calculate the voltage drop across the entire device, thus determining the current-voltage, I-V, characteristic.

6.2.2 Intrinsic region

In order to solve eqns. (6.1)-(6.3) semi-analytically some approximations are necessary. Two justifiable physical assumptions can be made inside the intrinsic region, [2]:

- *high-level* injection
  The injection of current from the two highly doped sides introduces a large amount of carriers inside the intrinsic region. Despite recombination processes, the total carrier density remains still high compared to the equilibrium values p_0, n_0. Therefore it is reasonable to assume

  $$p, n \gg p_0, n_0 \tag{6.7}$$

  At equilibrium, at room temperature, the intrinsic carrier density for Silicon is $p_0 = n_0 = 1.45 \cdot 10^{10} cm^{-3}$.

- *neutrality condition* or *charge balance*
  Inside the intrinsic region it is assumed that the excess electron density $\Delta n = n - n_0$ is exactly balanced by the excess hole density $\Delta p = p - p_0$, i.e,

  $$p - p_0 = \Delta p = \Delta n = n - n_0 \tag{6.8}$$

  This is equivalent of assuming charge neutrality, although it can not be exactly correct. In fact since the two types of carrier have different diffusion rate, $D_n \neq D_p$, the faster diffusing species (electron) would tend to diffuse more ahead than the slower diffusing species (hole), therefore separating altogether. Moreover, since carriers

are charged, their separation generates an electric field which tends to counteract the difference in speed for diffusion. Therefore, it is not be possible to achieve exactly the condition of charge neutrality. However, the difference between hole and electron densities is small compared with each carrier density, $p - n \ll p, n$, and the aforementioned assumption remains valid within good approximation.

Under the assumptions of *high-level* injection and *charge neutrality* it is possible to solve eqns. (6.1), (6.2) and find an analytic form for the carrier distribution, $p(x)$ and $n(x)$,

$$p(x) \approx n(x) = C_1 \cosh \frac{x}{L_i} + C_2 \sinh \frac{x}{L_i} \tag{6.9}$$

The proof of expression (6.9) and the values of the two constants, C_1, C_2, obtained by applying the boundary conditions (6.6) are detailed in Appendix E.

6.2.3 Current-Voltage (I-V) characteristic

The *current-voltage*, (I-V), characteristic of a p-i-n junction is similar to that of a p-n junction, although different in the slope: the difference becomes predominant for high current injection, which is the case analysed in this work.
Because the I-V curve shows, in general, a large slope it is convenient to treat the current rather than the voltage as the input parameter. In this way the numerical method is more stable since the most sensitive parameter (current) is used as input, and the voltage determined correspondingly: large variations of the current do indeed correspond to small variation of the voltage.
The analysis of the I-V characteristic is subdivided in two parts: inside the intrinsic region and inside the doped regions. The total voltage drop across the device is thus the summation of the three contributions

$$V_{pin} = V_{pi} + V_i + V_{in} \tag{6.10}$$

V_i is the voltage drop across the intrinsic region and V_{pi}, V_{in} are the voltage drops across the p-i and i-n junctions, respectively.
Since the electric field settles to a value which is almost zero after a short distance away from the junctions, inside the heavily doped regions, it is assumed that there is no voltage drop almost anywhere inside the doped regions, but only across some small distances near the junctions (depletion regions). Therefore, the total voltage drop V_{pin} is equal to the external voltage, V_{bias}, applied between the two metal contacts, Fig. [6.3].

Inside the intrinsic region the voltage drop is obtained by integration of the electric field profile, i.e.,

$$V_i = V(x = -l/2) - V(x = l/2) = -\int_{-l/2}^{l/2} [-E(x)] \mathrm{d}x \tag{6.11}$$

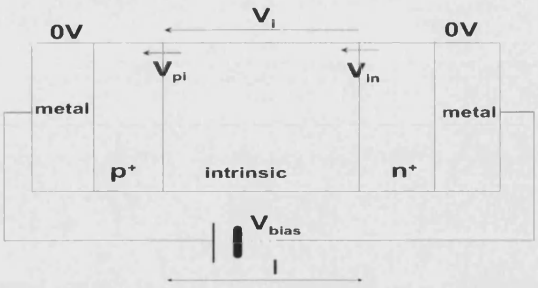


Figure 6.3: Voltage drop across the p-i-n structure.

The derivation of the integral (6.11) and the expression of V_{pi}, V_{in}, are described in Appendix F.

6.3 Two-dimensional analysis

In a more accurate two-dimensional analysis the current and the electric field vectors described in eqns. (5.2), (5.3), (5.10), must be considered, each constituted by two components, x and y as follows:

$$\begin{aligned}
\mathbf{J_p}(x,y) &= J_{px}(x,y)\hat{\mathbf{x}} + J_{py}(x,y)\hat{\mathbf{y}} \\
\mathbf{J_n}(x,y) &= J_{nx}(x,y)\hat{\mathbf{x}} + J_{ny}(x,y)\hat{\mathbf{y}} \\
\mathbf{E}(x,y) &= E_x(x,y)\hat{\mathbf{x}} + E_y(x,y)\hat{\mathbf{y}}
\end{aligned} \tag{6.12}$$

with $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ unity vectors. Moreover the device has been assumed uniform along the longitudinal z axis.
Combining eqns. (6.12) and (5.5), (5.6) the single components of the current are expressed in terms of carrier distributions:

$$\begin{aligned}
J_{px}(x,y) &= -eD_p\partial_x p(x,y) + \mu_p e p(x,y)E_x(x,y) \\
J_{py}(x,y) &= -eD_p\partial_y p(x,y) + \mu_p e p(x,y)E_y(x,y) \\
J_{nx}(x,y) &= eD_n\partial_x n(x,y) + \mu_n e n(x,y)E_x(x,y) \\
J_{ny}(x,y) &= eD_n\partial_y n(x,y) + \mu_n e n(x,y)E_y(x,y)
\end{aligned} \tag{6.13}$$

6.4 The Lumped Iterative Method (LIM)

In order to simplify the formalism the different types of recombination introduced in Section 5.2.1 are all represented through the lifetime constant, eqn. (5.12), so that the carrier conservation equations (5.2), (5.3), at the steady state, reduce to

$$\nabla \cdot \mathbf{J_p} = -e\frac{p(x,y)}{\tau_p} \tag{6.14}$$

$$\nabla \cdot \mathbf{J_n} = e\frac{n(x,y)}{\tau_n} \tag{6.15}$$

The basic strategy for solving the two-dimensional problem is the effective decoupling, with some approximations, of the lateral x profile and the transverse y profile of injected carriers. The structure in Fig. [6.1a] is considered as a lumped element, and described by two interconnecting wires, representing the horizontal and the vertical flux of carriers, Fig. [6.4]. The horizontal wire, representing the horizontal currents, J_{px}, J_{nx}, is connected to an applied external bias, while the flux of carriers along the rib, J_{py}, J_{ny}, is represented by a 'floating' wire, departing from the intersecting node with the horizontal wire. The node is the equivalent of the common central region A in Fig [6.1a].
In this way each direction is solved separately by reducing the two-dimensional solution to a one-dimensional problem, which can be easily solved in analytic form, as described in the previous section. The error introduced with the process of averaging is minimised by

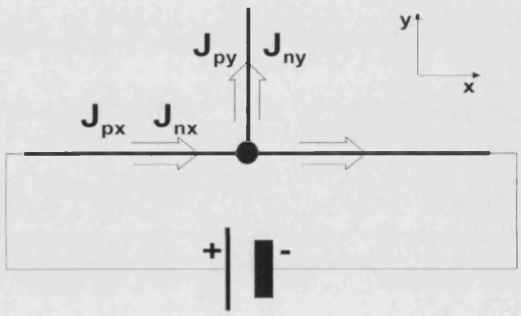


Figure 6.4: Lumped equivalent representation of the 2-D device.

applying an iterative method, until a convergent solution is achieved.
Each iteration cycle can be described by the following two steps, with the illustrative use of Fig. [6.5]:

1. In the first step the device is averaged in the y direction, so that structure in Fig. [6.1a] is reduced to a one-dimensional, (x-axis), horizontal device. The hole current conservation eqn. (6.14) is integrated in the y axis, between $y = -d$ and $y = 0$, Fig. [6.5a],

$$\frac{1}{d}\int_{-d}^{0}\left[\frac{\partial J_{px}(x,y)}{\partial x} + \frac{\partial J_{py}(x,y)}{\partial y}\right]\mathrm{d}y = \frac{-e}{d}\int_{-d}^{0}\frac{p(x,y)}{\tau_p}\mathrm{d}y \tag{6.16}$$

   The corresponding equation for electrons and electron current density is similar and is omitted here for brevity.
   Using the expression (5.5) for $\mathbf{J_p}$, and defining the quantity

$$\hat{p}(x) = \frac{1}{d}\int_{-d}^{0} p(x,y)\mathrm{d}y \tag{6.17}$$

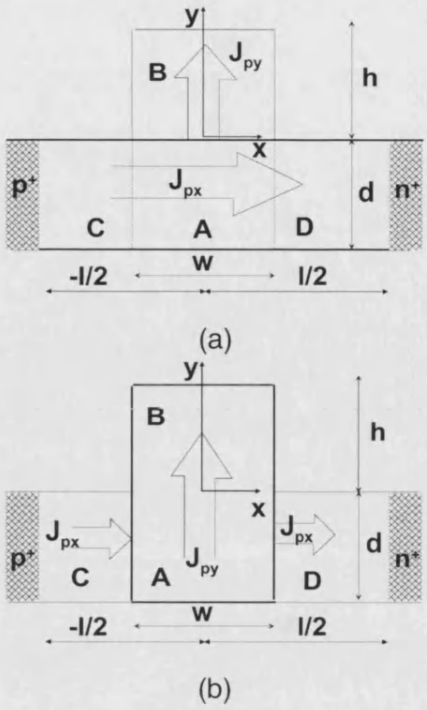


Figure 6.5: Averaging process. (a) device averaged in the y direction and solved along the x direction; (b) device averaged in the x direction and solved along the y direction.

eqn. (6.16) becomes

$$-D_p\frac{\mathrm{d}^2\hat{p}(x)}{\mathrm{d}x^2}+\mu_p\frac{\mathrm{d}}{\mathrm{d}x}[\hat{p}(x)\cdot\tilde{E}_x(x)]+\frac{\hat{p}(x)}{\tau_p}=\frac{1}{ed}[J_{py}(x,0)-J_{py}(x,-d)] \qquad (6.18)$$

In writing eqn. (6.18) it has been assumed that each component of the electric field, $\mathbf{E}(x,y)$, varies only along the corresponding direction, i.e. $E_x(x,y)\approx\tilde{E}_x(x)$, for the horizontal component, [3], and $E_y(x,y)\approx\tilde{E}_y(y)$ for the vertical component. In this way $\tilde{E}_x(x)$ acts as a constant in averaging over the cross section at fixed x, and it is in accordance with $\nabla\times\mathbf{E}=\mathbf{0}$ (there is no time varying field in the problem), i.e. $\partial_x E_y-\partial_y E_x=0$, [3].

The physical meaning of eqn. (6.18) can be visualised in Fig. [6.5a]; the analysis is focused on the horizontal flux of current, J_{px}, inside the region delineated by the surfaces $y=-d$ and $y=0$. Following the averaged representation, eqn. (6.17), this region is considered uniform along the vertical y direction. However, the effects of the top rib are represented through the current J_{py} escaping out of the horizontal region, Fig. [6.5a].

In the first iteration the rib is neglected, thus

$$J_{py}(x,-d)=0 \quad J_{py}(x,0)=0 \quad \text{for all values of x} \qquad (6.19)$$

since carriers do not flow out of material surfaces (the case of surface recombination is considered in detail at a later stage in the thesis). As a consequence, at the first

iteration, eqn. (6.18) reduces to the same form of the 1-D case, eqn. (6.1),

$$-D_p\frac{\mathrm{d}^2\hat{p}(x)}{\mathrm{d}x^2}+\mu_p\frac{\mathrm{d}}{\mathrm{d}x}[\hat{p}(x)\cdot\tilde{E}_x(x)]+\frac{\hat{p}(x)}{\tau_p}=0 \tag{6.20}$$

Eqn. (6.20) together with the analogous equation for electron carriers can be treated equivalently to the one-dimensional eqns. (6.1), (6.2), obtaining an analytic solution for the horizontal carrier profiles $\hat{p}(x)$. However, boundary conditions (6.6) can not be assigned explicitly to $\hat{p}(x)$, unless a separation of variables is assumed, i.e.,

$$p(x,y)=f(x)\cdot g(y) \qquad n(x,y)=u(x)\cdot v(y) \tag{6.21}$$

With this choice eqn. (6.17) becomes

$$\hat{p}(x)=f(x)\cdot\frac{1}{d}\int_{-d}^{0}g(y)\mathrm{d}y=f(x)\cdot\hat{g} \tag{6.22}$$

and an analytic solution for eqn. (6.20) is obtained

$$\hat{p}(x)=\left(a_1\cosh\frac{x}{L_i}+a_2\sinh\frac{x}{L_i}\right)\cdot\hat{g} \tag{6.23}$$

where the constants a_1, a_2, are determined by applying the same boundary conditions as in eqn. (6.6),

$$J_{nx}(x=-l/2)=0 \qquad J_{px}(x=l/2)=0 \tag{6.24}$$

In successive iterations the effect of the rib on the horizontal flux of carriers is reintroduced, considering the value of the drain current, J_{py} in eqn. (6.18). At the bottom interface, $y=-d$, there is still no current flow, while at the top interface, $y=0$, there is a drain of carriers toward the rib region, J_{py} corresponding to the position of the rib, $|x|<w/2$. The value of J_{py} is estimated by using the results of the vertical current flow obtained in step (2)

$$\begin{aligned} &J_{py}(x,-d)=0 \quad \text{for all values of x}\\ &J_{py}(x,0)=\begin{cases} J_{pv}(y=0) & \text{for } |x|<w/2\\ 0 & \text{for } |x|>w/2 \end{cases} \end{aligned} \tag{6.25}$$

With this choice eqn. (6.18), in the range $|x|<w/2$, becomes

$$-D_p\frac{\mathrm{d}^2\hat{p}(x)}{\mathrm{d}x^2}+\mu_p\frac{\mathrm{d}}{\mathrm{d}x}[\hat{p}(x)\cdot\tilde{E}_x(x)]+\frac{\hat{p}(x)}{\tau_p}=\frac{1}{ed}J_{pv}(y=0) \tag{6.26}$$

while in the range $|x|>w/2$ it has the same form of eqn. (6.20), and similarly for electrons. The solution is uniquely determined by applying the boundary conditions (6.24) at the two boundaries, $x=\pm l/2$, and requiring the continuity of $\hat{p}(x)$ at the common interfaces, $x=\pm w/2$.

2. In the second-step the hole carrier eqn. (6.14) is solved along the vertical y direction. The problem is equivalent of solving a one-dimensional structure as in step (1), except that now carrier profiles are solved without any x dependence. This is achieved by averaging eqns. (6.14) along x. Thus, referring to Fig. [6.5b], it is

$$\frac{1}{w}\int_{-w/2}^{w/2}\left[\frac{\partial J_{px}(x,y)}{\partial x}+\frac{\partial J_{py}(x,y)}{\partial y}\right]\mathrm{d}x=-\frac{e}{w}\int_{-w/2}^{w/2}\frac{p(x,y)}{\tau_p}\mathrm{d}x \tag{6.27}$$

and similarly for electrons.

With analogy to step (1) the horizontal average value of the carrier profile across the central regions A and B, Fig. [6.5b], is defined as

$$\bar{p}(y)=\frac{1}{w}\int_{-w/2}^{w/2}p(x,y)\mathrm{d}x \tag{6.28}$$

Therefore eqn. (6.27), with the use of eqn. (5.5) and the assumption on the electric field $E_y(x,y)\approx\tilde{E}_y(y)$, reduces to

$$-D_p\frac{\mathrm{d}^2\bar{p}(y)}{\mathrm{d}y^2}+\mu_p\frac{\mathrm{d}}{\mathrm{d}y}[\bar{p}(y)\cdot\tilde{E}_y(y)]+\frac{\bar{p}(y)}{\tau_p}=\frac{1}{ew}[J_{px}(-w/2,y)-J_{px}(w/2,y)] \tag{6.29}$$

The illustrative explanation of eqn. (6.29) is given in Fig. [6.5b]. The current J_{px} in eqn. (6.29) accounts for the escape of carriers out of regions A and B, at the two interfaces $x=\pm w/2$ and it is non-zero only in the range $-d<y<0$, because carriers can not escape outside the boundaries of the rib region, for $y>0$. The value of $J_{px}(\pm w/2,y)$ is approximated with the value of J_{px} obtained in step (1),

$$J_{px}(\pm w/2,y)=\begin{cases}J_{ph}(\pm w/2) & \text{for } -d<y<0\\ 0 & \text{for } y>0\end{cases} \tag{6.30}$$

The solution is determined by satisfying the continuity of $\bar{p}(y)$ at the common interface, $y=0$, and applying the boundary conditions at the physical interfaces, $y=-d$ and $y=h$. After invoking the separation of variable (6.21), it is

$$\bar{p}(y)=\frac{1}{w}\int_{-w/2}^{w/2}f(x)\mathrm{d}x\cdot g(y)=\bar{f}\cdot g(y) \tag{6.31}$$

and the following boundary conditions are applied to determine $\bar{p}(y)$ uniquely:

$$J_{py}(y=h)=0 \qquad J_{py}(y=-d)=0 \tag{6.32}$$

considering that carriers can not flow out through the top and bottom edges of the rib, respectively.

Steps (1) and (2) are iterated, thus refining the solution for the carrier profiles, $p(x,y)$, at each iteration, until the solution converges to a steady distribution.

It is useful to note that from a circuit analysis the vertical arm in the lumped circuit, Fig.(6.4), is an open circuit. Therefore the total amount of current coming out of it is zero, $J_y = J_{py} + J_{ny} = 0$. This condition is true for any value $y > 0$, along the vertical wire. In particular, it is true at the top edge of the rib, $y = h$, where both electron and hole currents are separately zero, i.e., $J_{py} = 0$, $J_{ny} = 0$. However, in general this does not imply that hole and electron currents are each separately zero everywhere: for example, corresponding to the node, at $y = 0$, the two vertical carrier currents have a nonzero value, $J_{py}(y = 0) \neq 0$, $J_{ny}(y = 0) \neq 0$.

6.4.1 Surface Recombination

A further refinement of the model can be accomplished by introducing the effects of surface recombination. As introduced in Section 5.2.1 surface recombination can not be modelled through an effective carrier lifetime, but must be considered as a boundary condition at the pertinent surface. The result of such recombination is that holes and electrons will recombine at the surface, generating an effective flux of carriers through the surface, [4]. It is important to note that, because one electron recombines with one hole, the flux of holes at the surface is precisely equal to the flux of electrons, thus resulting in no net current flow.

In order to model surface recombination in the Lumped Iterative Method, the conditions on the flux of current at the surfaces must be revised. In particular, in the first step it is assumed that there is a vertical flux of current, $J_{py}(x, 0)$, at the interface $y = 0$, even corresponding to the regions C and D, Fig. [6.6a]; thus rewriting eqn. (6.25) as

$$\begin{aligned} J_{py}(x, -d) &= 0 \quad \text{for all values of x} \\ J_{py}(x, 0) &= \begin{cases} J_{pv}(y = 0) & \text{for } |x| < w/2 \\ eS_r p(x, 0) & \text{for } |x| > w/2 \end{cases} \end{aligned} \tag{6.33}$$

with S_r the surface recombination velocity, [4].

Similarly, in step (2) the current $J_{px}(\pm w/2, y)$ in eqn. (6.30) is replaced by, Fig. [6.6b],

$$J_{px}(\pm w/2, y) = \begin{cases} J_{ph}(\pm w/2) & \text{for } -d < y < 0 \\ \pm eS_r p(\pm w/2, y) & \text{for } y > 0 \end{cases} \tag{6.34}$$

Moreover, in the second step also the boundary condition (6.32) at the top edge of the rib, $y = h$, is revised and surface recombination included, i.e.,

$$J_{py}(y = h) = eS_r p(x, h) \tag{6.35}$$

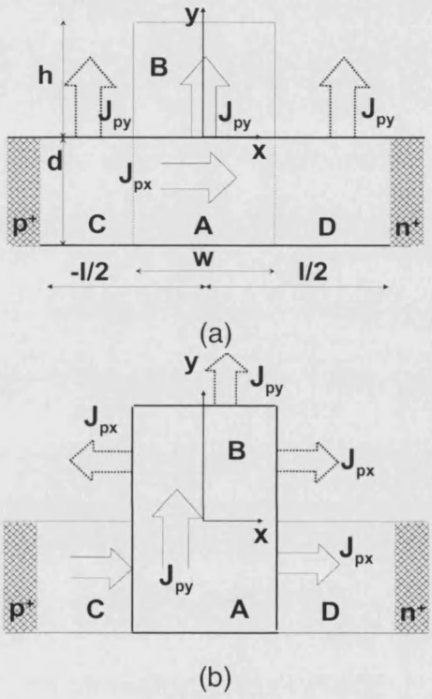


Figure 6.6: Averaging process with surface recombinations. The effect of surface recombinations is represented by the extra drain of carriers at the surface (dotted flux of current in the picture). (a) device averaged in the y direction and solved along the x direction; (b) device averaged in the x direction and solved along the y direction.

Bibliography

[1] Gummel H.K., "A Self-Consistent Iterative Scheme for One-Dimensional Steady State Transistor Calculations", *IEEE Trans. on Elec. Dev.*, 1964, 455-465

[2] McKelvey J.P., "Solid State and Semiconductor Physics", Harper & Row, NY, 1966

[3] Lampert M.A., Mark P., "Current Injection in Solids", Academic Press, NY, 1970

[4] Grove A.S., "Physic and Technology of Semiconductor Devices", Wiley, 1967

Chapter 7

Results: carrier distribution

In this chapter some numerical results for the carrier distributions obtained by applying the Lumped Iterative Method methods (LIM) are presented.
The structure analysed is depicted in Fig. [7.1]. Dimensions and semiconductor paramet-

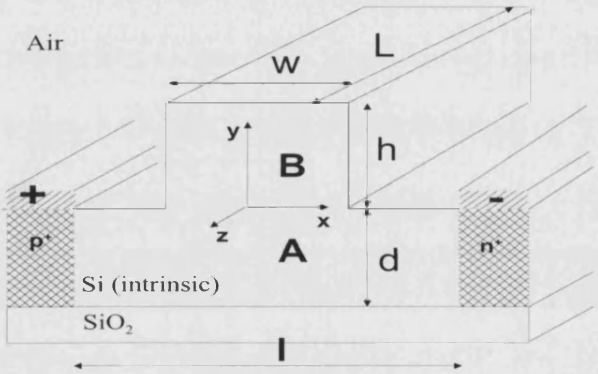


Figure 7.1: Structure of the rib waveguide analysed.

ers utilised for the simulation are as in Tab. [7.1]. The LIM determines the two-dimensional carrier distributions across the transverse x-y plane. The longitudinal length, L, of the device, Fig. [7.1], is reported only to provide the link between current densities and injected currents. In general, the device can be of any length, L. The particular length, $L = 2000\mu m$, chosen in Tab. [7.1] reflects typical values of Silicon-on-Insulator realistic devices.

The carrier distribution is firstly calculated for a one-dimensional p-i-n junction, inside both the doped and the intrinsic regions. Then the LIM is applied to determine the two-dimensional profile of carrier and current flow distributions. Results are compared with those obtained by using a Finite Element Method, showing good agreement. The effect of the rib on the carrier profile is discussed in details, together with the influence of recombination processes. Finally, the efficacy of the LIM as a way to predict the refractive index change of the material, prior to carrier injection, is demonstrated by comparing the results

dimensions:
$w = 3.5\mu m$, $d = 2.6\mu m$, $h = 1.6\mu m$, $l = 12\mu m$, $L = 2000\mu m$
recombination coefficients:
$\tau_{p0} = \tau_{n0} = 2 \cdot 10^{-7} s$
$C_p = 1.8 \cdot 10^{-31} cm^6/s$, $C_n = 8.3 \cdot 10^{-32} cm^6/s$
$S_r = 8 \cdot 10^4 cm/s$

Table 7.1: Dimensions of the rib waveguide and recombination coefficients for Silicon material.

with a similar published work.

Reference model

In order to test the accuracy of the Lumped Iterative Method the commercial software Silvaco has been adopted as a reference model. Silvaco is designed to solve the continuity flow eqns. (5.2), (5.3), together with Poisson's eqn. (5.10) by using a Finite Element numerical technique.

In general, the carrier mobility, μ_p, μ_n, in semiconductor materials depend on the carrier concentrations, [1]. This effect is not directly taken into account by the LIM; however, at each step of iteration, μ_p and μ_n are considered constant, and their values are set according to the hole and electron concentrations obtained as results in the previous step. These values are tabulated in literature: in the particular case of Silicon the mobility dependence of the carrier concentration is shown in Fig. [7.2], [2]. Because most of the devices under

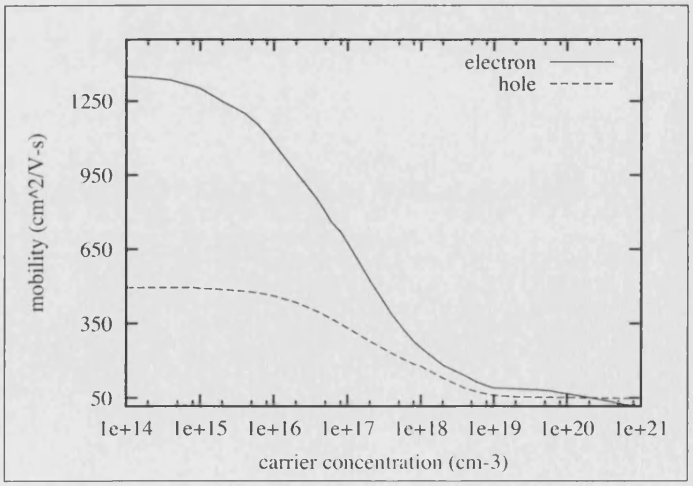


Figure 7.2: Mobilities dependence (experimental) of the carrier concentration, N (in cm^{-3}). Solid line is $\mu_n(N)$; dashed line is $\mu_p(N)$, in $(cm^2/V - s)$., [2]

investigation operate in the range of *high-level* injection it is possible to assume, by direct observation of Fig.[7.2], that the values of μ_p and μ_n are almost constant. Therefore, as a first approximation it may also be possible to consider the mobilities as constants throughout the entire iterative process.

7.1 One-dimensional p-i-n junction

As a first approximation the device in Fig. [7.1] is solved as a one-dimensional p-i-n junction, along the x direction, neglecting the presence of the rib.
The carrier concentrations are analysed according to the model described in Section 6.2 for the one-dimensional p-i-n junction. The computation is repeated for two different cases of recombination: (a) considering the Shockley-Read-Moll (SRM) non-radiative recombination only; (b) considering both SRM and Auger recombination.
The amount of injected current is varied between $I_0 = 1mA$ and $I_0 = 100mA$. Typical profiles of the hole carrier concentration inside the intrinsic region are shown in Fig. [7.3], for different values of injected current. The hole and electron concentrations are almost the same, $p(x) \approx n(x)$, since inside the intrinsic region the neutrality condition, eqn. (6.8), is valid.
Results are compared, Tab. [7.2], with those obtained by applying the reference software

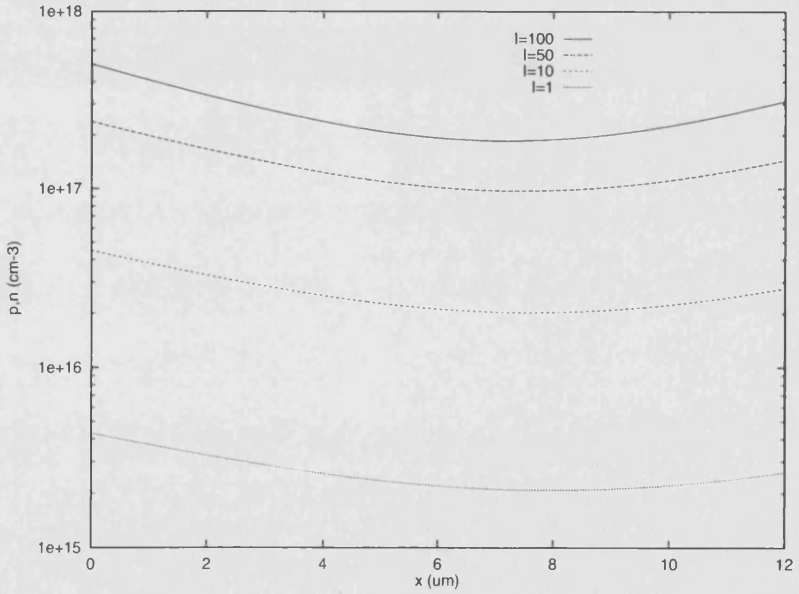


Figure 7.3: Carrier density distribution, $p(x)$ (in cm^{-3}), inside the intrinsic region, for different values of injected current. $I_0 = 1, 10, 50, 100mA$. Only the SRM recombination has been included.

Silvaco, which accounts for both SRM and Auger recombination effects.
All results are in agreement for small values of injected current. Since cases (a) and (b)

current I_0 (mA)	Silvaco	case(a)	case (b)
1	$4 \cdot 10^{15}$	$4 \cdot 10^{15}$	$4 \cdot 10^{15}$
10	$3 \cdot 10^{16}$	$4 \cdot 10^{16}$	$4 \cdot 10^{16}$
50	10^{17}	$2 \cdot 10^{17}$	$1.5 \cdot 10^{17}$
100	$1.5 \cdot 10^{17}$	$4 \cdot 10^{17}$	$2 \cdot 10^{17}$

Table 7.2: Comparison of carrier density (in cm^{-3}), at $x = 0$, between Silvaco and the present method. Case (a): SRM recombination only is considered; case (b): SRM and Auger recombinations are considered.

provide similar results Auger recombination is negligible. Conversely, with large current injected Auger effect becomes predominant. In fact, if Auger recombination is omitted, as in case (a), i.e. the total recombination process is underestimated, the LIM provides values of carrier concentration divergent from those obtained by Silvaco.
It is important to note how the assumption of *high-level* injection is valid even in the case of small values of injected current, since the amount of injected carriers is still much larger than the equilibrium concentration, $p_0 = n_0 = 1.45 \cdot 10^{10} cm^{-3}$.

Inside the p doped region, far apart from the p-i junction, the charge neutrality condition is still valid. Since this region is doped, at equilibrium it is $p_0 \gg n_0$. Therefore, in order to satisfy the charge neutrality condition, eqn. (6.8), it is necessary that $p \gg n$ and $p \approx p_0$. It is then expected that, even at non-equilibrium, the hole density settles to a value near the equilibrium concentration, p_0.
The transition of the hole concentration from the heavily p doped region, $p \approx p_0$, to the value at the beginning of the intrinsic region, $p(x = -l/2)$, Fig. [7.1], occurs within few Angstroms (depletion region), Fig. [7.4]. Inside the transition region the charge neutrality condition is not valid, and a high electric field is generated.
It is possible to estimate the length of this transition region from conventional p-n junction theory. In general the depletion region, where most of the transition occurs, for a standard semiconductor material is long $T_p = sL_d$, Fig. [6.2], where s is a constant depending of the material, and L_d is the Debye length, defined as, [3],

$$L_d = \sqrt{\frac{\epsilon V_t}{eN_a}} \tag{7.1}$$

For Silicon material $T_p \approx 8L_d$. The Debye length, L_d, is proportional to the doping density as $1/\sqrt{N_a}$. With the actual doping density of $N_a = 10^{20} cm^{-3}$, the corresponding Debye length is $L_d = 4.15$Å, hence, the transition region is approximately $T_p = 33$Å. This value is in agreement with the results obtained in Fig. [7.4], where the hole concentration reaches its equilibrium value, $p = p_0 = N_a = 10^{20} cm^{-3}$, after a distance of approximately $8L_d$ from

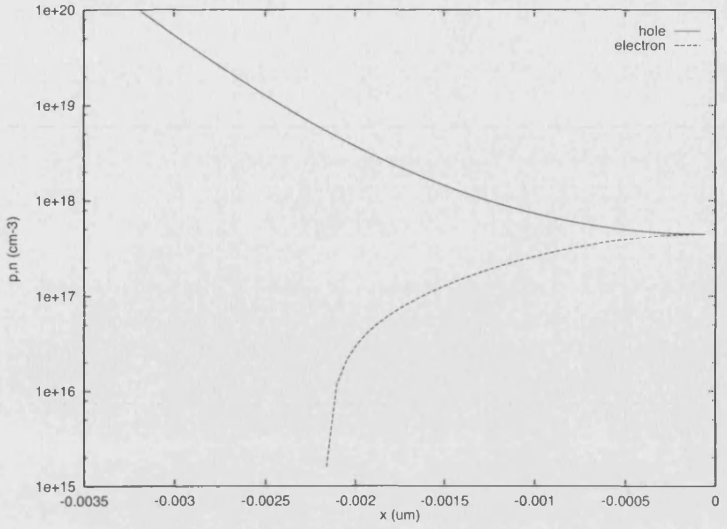


Figure 7.4: Carrier distributions inside the p doped region. Solid line is $p(x)$, dashed line is $n(x)$ (in cm^{-3}). Doping density is $N_a = 10^{20} cm^{-3}$.

the junction, into the p side doped region.
In lower doped material the carrier profiles reach the equilibrium value at a longer distance from the junction, Fig. [7.5]. Carrier profiles at the i-n junction can be explained in a similar way, with electrons replacing holes.

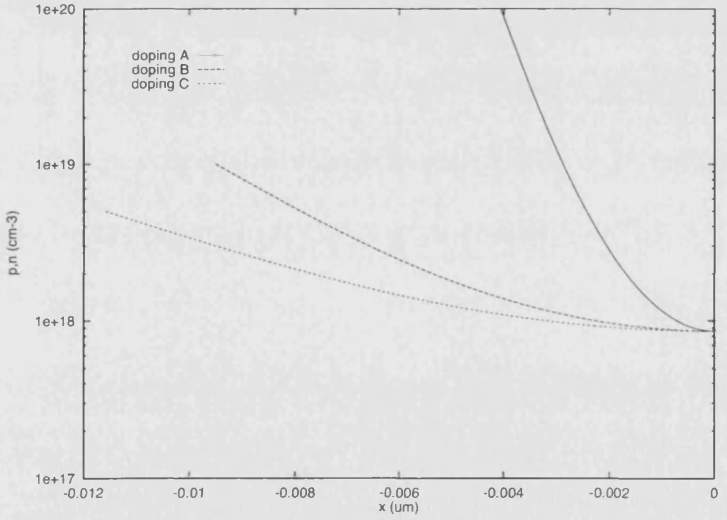


Figure 7.5: Majority carrier distribution in the p doped region, for different amount of doping density, N_a (in cm^{-3}). A: $N_a = 10^{20}$; B: $N_a = 10^{19}$; C: $N_a = 10^{18}$.

7.1.1 I-V characteristic and Current densities

As described in Section 6.2.3 once the carrier profiles are established it is possible to determine the total voltage drop across the p-i-n junction. Moreover, by iterating the calculation for different values of injected current, and obtaining the correspondent voltage drop, it is possible to characterise the current-voltage (I-V) curve.
As an example, Fig. [7.6] shows the I-V characteristic for a p-i-n junction with dimensions and parameters as indicated in Tab. [7.1].
For completeness, the profile of the total current across the intrinsic region is documented,

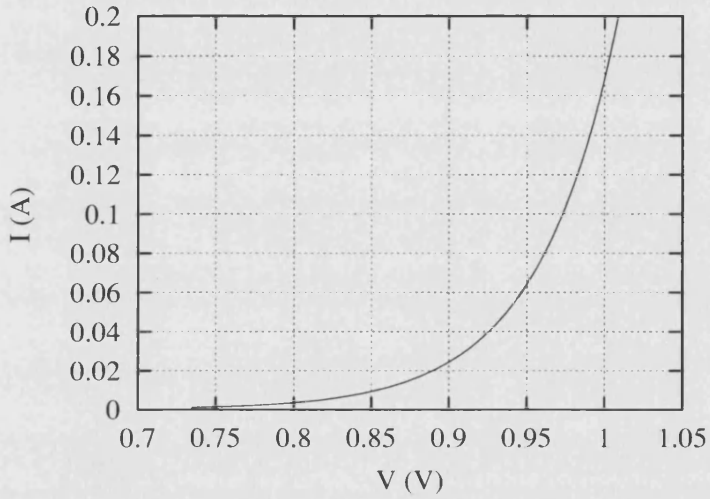


Figure 7.6: Current-Voltage characteristic for a p-i-n junction device.

for an injected current of $I_0 = 100mA$. Both hole and electron currents are determined from the carrier profiles, and shown in Fig. [7.7] and Fig. [7.8], respectively. In accordance with the boundary conditions (6.5) there is no flux of electron at the p-i junction ($x = -l/2$) and no flux of holes at the i-n junction ($x = l/2$). Moreover the total current density, $J_{tot} = J_{ptot} + J_{ntot}$, as shown in Fig. [7.9], is constant across the device.

7.2 2-D Carrier analysis: the Lumped Iterative Method

The two-dimensional carrier distribution for the device depicted in Fig. [7.1] is analysed by applying the Lumped Iterative Method (LIM). Dimensions and parameters utilised in the simulation are the same as those described in Tab. [7.1]. As first instance surface recombinations are neglected: the effects of surfaces will be re-introduced and emphasised at a later stage in the chapter.
A typical two-dimensional hole density profile across the intrinsic region, obtained by applying the LIM, is shown in Fig. [7.10]. Because the analysis is focused on the intrinsic

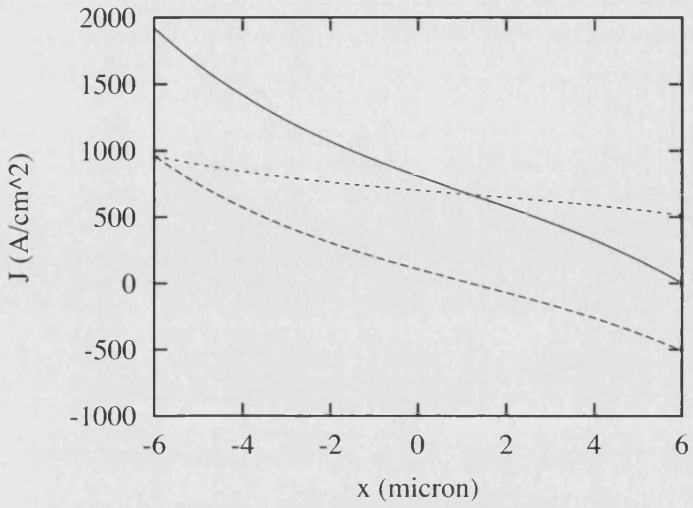


Figure 7.7: Hole current density (in A/cm^2). Solid line: total hole current density, J_{ptot}; Dotted line: conduction current density; Dashed line: diffusion current density.

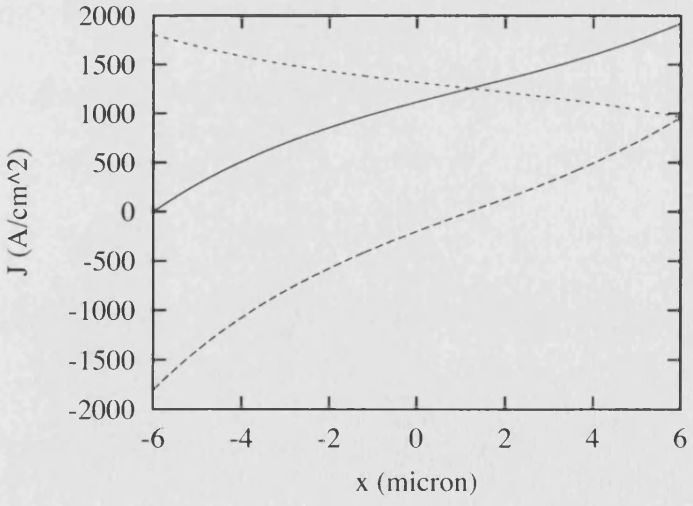


Figure 7.8: Electron current density (in A/cm^2). Solid line: total hole current density, J_{ntot}; Dotted line: conduction current density; Dashed line: diffusion current density.

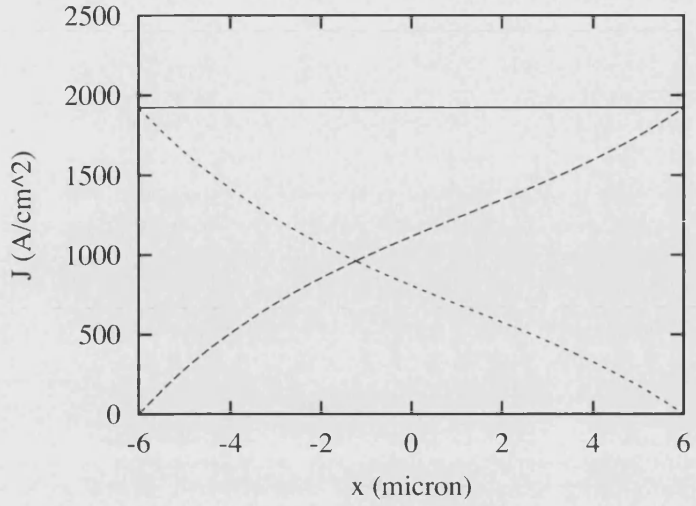


Figure 7.9: Current density across the intrinsic region (in A/cm^2). Solid line: total current density, J_{tot}; Dotted line: hole current density, J_p; Dashed line: electron current density, J_n.

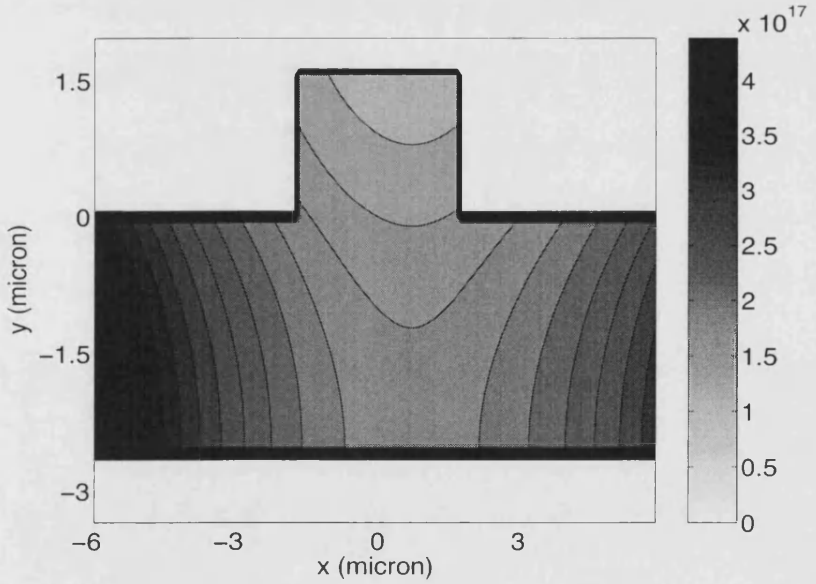


Figure 7.10: Contour plot of the carrier density profile, $p(x,y) = n(x,y)$ (in cm^{-3}), inside the intrinsic region. Injected current: I=50mA.

region where neutrality has been assumed, the two types of carriers, hole and electron, have almost the same value.
Part of the carriers redistribute also inside the rib region, hence reducing the amount of carriers in the region A underneath the rib, Fig. [7.1]. Thus the rib does have an effect on the carrier distribution and, therefore, the analysis conducted by approximating the device with a simple p-i-n junction may not be accurate and sufficient to account for all these effects.

Effect of the rib on the carrier distribution

The influence of the rib can be estimated by comparing the profiles of the carrier concentration for devices with different rib widths. Fig. [7.11] depicts the carrier profile taken at the horizontal cross section below the rib, correspondingly to $y = -1\mu m$, Fig. [7.1]. In the case of the wider rib a larger amount of carriers flow inside the rib region, so that a smaller amount of carrier remains below the rib (the injected current is kept constant). The results presented in Fig. [7.11] are also compared with those obtained by utilising the commercial software Silvaco showing good agreement for the levels of the carrier concentration.

Same effects can be visualised from a different point of view by observing the vertical cross-section of the 2-D carrier profile at $x = 0$, Fig. [7.12]. The device with the larger stripe ($w = 7\mu m$) forces a larger drain of carriers inside the rib region. However, since the amount of injected carriers is the same the extra flux of carriers toward the rib does reduce more the amount of carriers left below the rib ($y < 0$).

A further test has been conducted by comparing the vertical carrier profile of two devices with same rib width, w, but different rib height, h. Fig. [7.13] depicts how carriers redistribute more in the device with the higher rib ($h = 3.5\mu m$) than in the case of the shorter rib ($h = 1.6\mu m$), thus reducing the level of concentration in the central region below the rib ($y < 0$).

The limit case has also been analysed by applying the LIM to a large rib, approaching a 1-D p-i-n junction. A rib with width $w = l$ has been considered, so that the carrier distribution is similar to the 1-D case where there is no rib, $h = 0$, but the height d of the bottom region is increased, Fig. [7.1], as such that the total vertical height $d + h$ is kept the same in both cases, Fig. [7.14].

7.2.1 Current distribution

The distribution of carriers, due either to diffusion or to conduction, does generate a flux of current across the device. Because the 2-D carrier profile is not uniform both the horizontal and the vertical components of the current are generated, i.e. all the four current components in eqn. (6.13) are present.

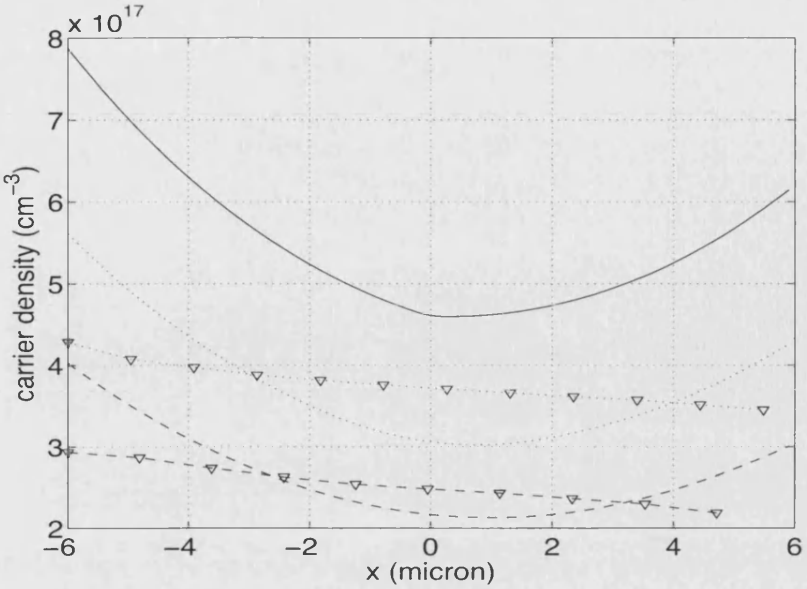


Figure 7.11: Carrier profile $p(x, y)$ (in cm^{-3}) at the horizontal cross-section $y = -1\mu m$, Fig. [7.1]. The curved solid line corresponds to a uniform device without rib; the curved dotted and dashed lines refer to rib devices with rib width of $w = 3.5\mu m$ and $w = 7\mu m$, respectively, for the same amount of injected current, $I_0 = 50mA$. The dotted-diamond and dashed-diamond line represents results obtained by utilising Silvaco to solve the same devices ($w = 3.5\mu m$ and $w = 7\mu m$), respectively. Injected current: $I_0 = 50mA$.

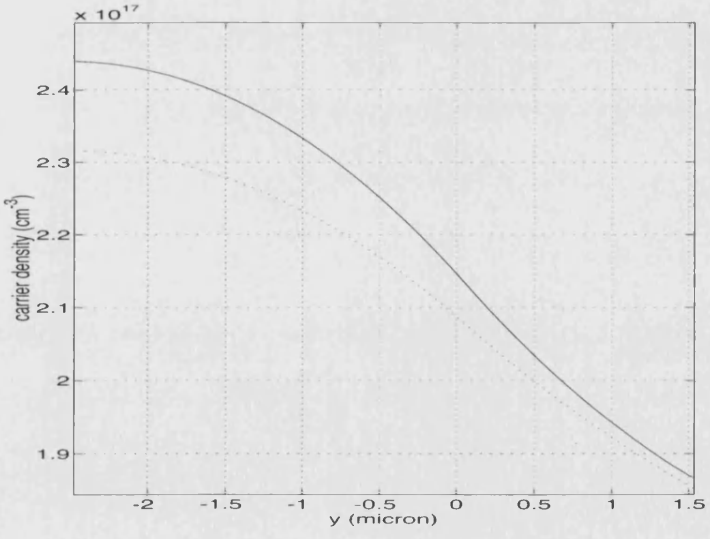


Figure 7.12: Carrier profile, $p(x, y)$ (in cm^{-3}), at the vertical cross-section $x = 0$, Fig. [7.1]. Solid line: device with rib width $w = 3.5\mu m$; Dotted line: device with rib width $w = 7\mu m$. Injected current: $I_0 = 50mA$.

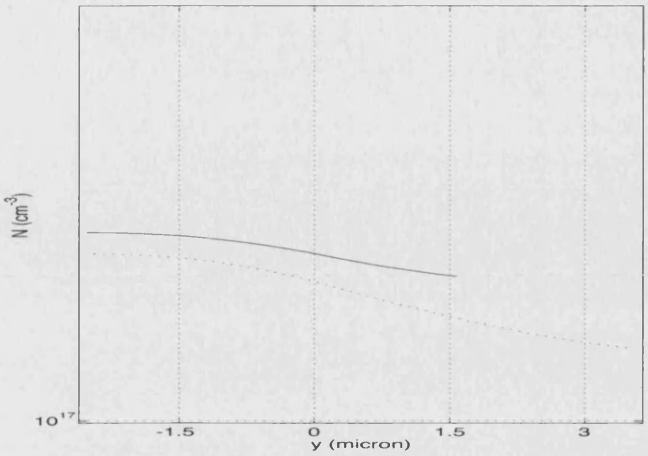


Figure 7.13: Carrier profile, $p(x, y)$ (in cm^{-3}), at the vertical cross-section $x = 0$, Fig. [7.1]. Solid line: device with rib height $h = 1.6\mu m$; Dotted line: device with rib height $h = 3.5\mu m$. The rib region corresponds to the region $y > 0$. Injected current: $I_0 = 50mA$.

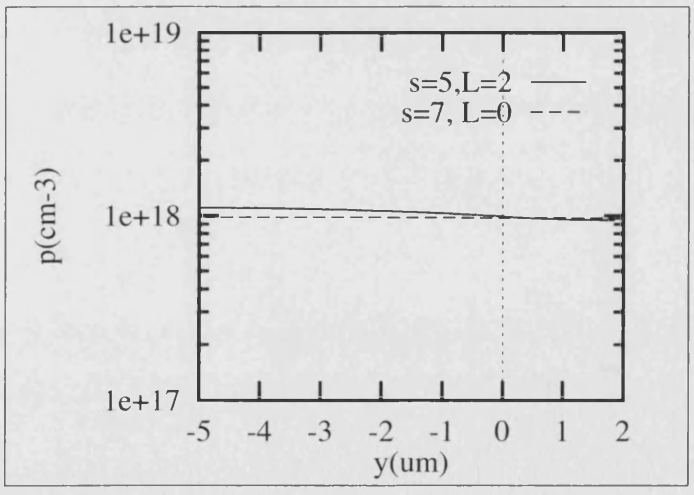


Figure 7.14: Carrier profile $p(x, y)$ (in cm^{-3}), at the cross section $x = 0$. The total height is kept constant ($d+h = 7\mu m$), the rib width is $w = l$, Fig. [7.1]: Solid line $d = 5\mu m$, $h = 2\mu m$; Dashed line: and $d = 7\mu m$, $h = 0\mu m$. Injected current: $I_0 = 50mA$.

Fig. [7.15] shows both hole and electron current flows. Accordingly to the boundary conditions (6.5) there is no hole current flowing at the i-n junction, ($x = l/2$), and there is no electron current flowing at the p-i junction, ($x = -l/2$), Fig. [7.1]. The major flux of current occurs between the two sides of the device, where the electrodes are positioned. However, in Fig. [7.15] it is also possible to observe flux of currents, J_p, J_n, toward the rib region, where there are no electrodes. These fluxes are generated by the gradient of carrier concentrations between the bottom region and the rib (diffusion currents). In addition, part of the vertical flux is generated by surface recombinations at the walls of the rib. Despite each type of carrier (hole and electron) generate a diffusion current, the total current, $J_{tot} = J_p + J_n$, is zero, so that, as expected, there is no net flux of current inside the rib.

The current flows from the p side to the n side since the device is forward biased; however, the flux is not uniform along the vertical direction, as in proximity of the rib it is influenced and partially distorted inside the rib. This feature is visualised in Fig. [7.16] where the horizontal component of hole, J_{px}, electron, J_{nx}, and total current, $J_x = J_{px} + J_{nx}$, are reported. The assumptions made in the Lumped Iterative Method implies that the net vertical flux of current inside the rib is zero, i.e. $J_{py} + J_{ny} = 0$, as it corresponds to the current flowing through the arm of an open circuit, Fig. [6.4]. However, this does not imply that both hole and electron currents must be exactly zero, except at the top of the rib. This concept is visualised in Fig. [7.17]: each carrier current is different from zero, although there is no net flux of current. However, corresponding to the top of the rib ($y = 1.6\mu m$) both electron and hole currents have zero value.

In both Figs. [7.16]-[7.17] the results are also compared with those obtained by applying the reference software Silvaco, showing similar profiles.

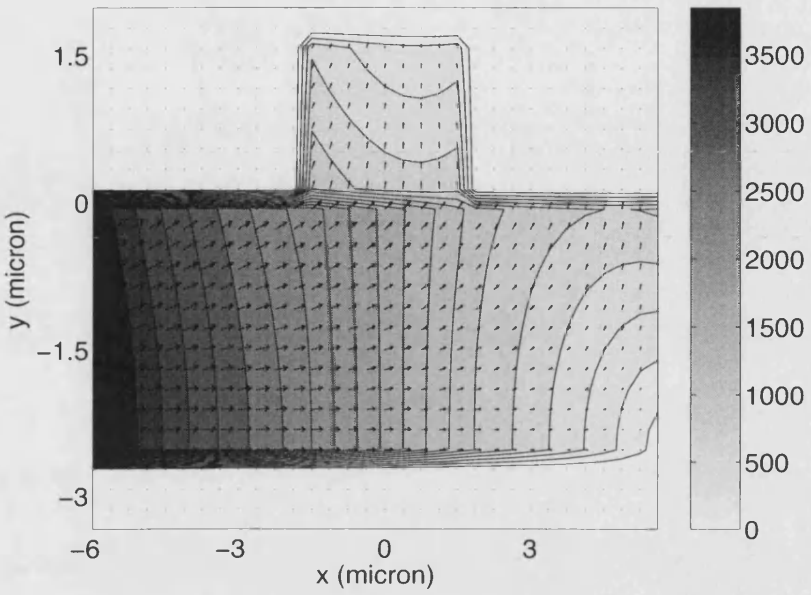
7.2.2 Effects of recombination processes

Recombination of carriers can be classified as: (a) bulk recombination, i.e. Shockley-Read-Moll (SRM), Auger; (b) surface recombination.

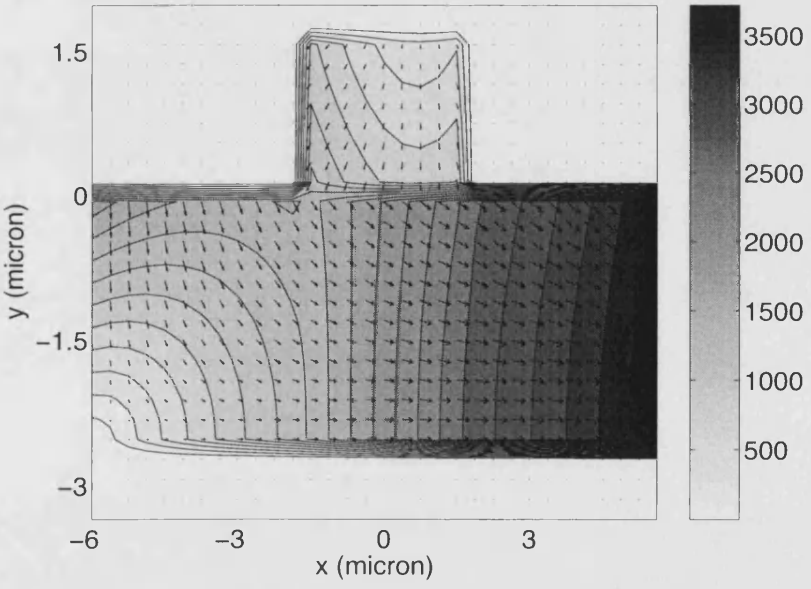
In case (a) recombinations are distributed across the whole device and represented in the model through the effective lifetime. A shorter lifetime does determine carriers to recombine at a faster rate.

Fig. [7.18] focuses on the effects of the non-radiative traps SRM recombination by comparing the carrier profile for two different values of carrier lifetime, eqn. (5.17), $\tau_{p0} = \tau_{n0} = 2 \cdot 10^{-7} s$ and $\tau_{p0} = \tau_{n0} = 2 \cdot 10^{-9} s$. In the case of short lifetime more carriers would recombine by the time they reach the central part of the intrinsic region and the level of carriers in the central region does show a deeper well. The results are in agreement with those obtained by utilising Silvaco.

Differently from bulk recombination surface recombination occurs predominantly in prox-



(a)



(b)

Figure 7.15: Current flow across the device (in A/cm^2). (a) Hole current density, $\mathbf{J_p} = (J_{px}, J_{py})$; (b) Electron current density, $\mathbf{J_n} = (J_{nx}, J_{ny})$. Injected current: $I_0 = 200mA$.

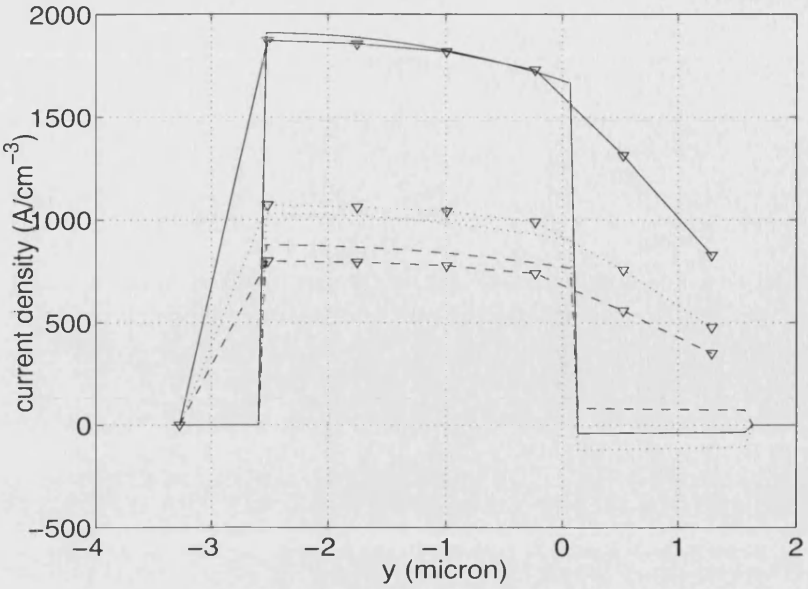


Figure 7.16: Profiles of the horizontal component of the current density (in A/cm^2) at the cross section $x = 0$, Fig [7.1]. Solid line: total current density, $J_x = J_{px} + J_{nx}$; Dotted line: electron current density, J_{nx}; Dashed line: hole current density, J_{px}; Solid-doamond, dotted-doamond, dashed-diamond: same as before, but results obtained by applying the software Silvaco. Injected current: $I_0 = 200mA$..

imity of surfaces. The Lumped Iterative Method does take in to account surface recombinations and the effects are shown in Fig. [7.19]. The surface recombination velocity, S_r, for Silicon/Air interface lies between $10^2 cm/s$ and $8 \cdot 10^4 cm/s$, while at the bottom Si/SiO_2 interface, Fig. [7.1], can be as small as $0.5cm/s$, [4]. Therefore it is sufficient to investigate the effect of surface recombinations only at the top Si/Air surface.

Fig. [7.19] shows how the carrier profile is strongly affected by the top Si/Air surface, when surface recombinations are considered. On the other hand, at a distance from such surface, in the bottom part of the device, the vertical gradient of the carrier distribution remains unaffected, independently of the top surface recombination velocity, S_r, that is considered.

7.2.3 Accuracy of the Lumped Iterative Method

The validity of the Lumped Iterative Method as an accurate way to determine the carrier concentration (and the consequent induced refractive index change of the material) has been tested by comparing the results with a similar published work. Hewitt et al. in [5] performed a more systematic application of the software Silvaco to determine the refractive index change, after injecting a current into a semiconductor rib device.
The parameters for the structure analysed in [5] with reference to Fig. [7.1] are indicated

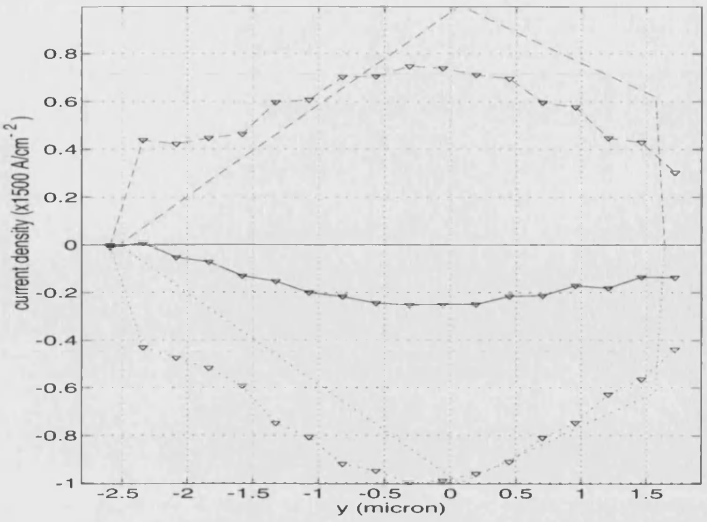


Figure 7.17: Vertical flux of current (in A/cm^2) at the cross-section $x = 0$, Fig [7.1]. Solid line: total current density, $J_y = J_{py} + J_{ny}$; Dotted line: electron current density, J_{ny}; Dashed line: hole current density, J_{py}; Solid-diamond, dotted-diamond, dashed-diamond: same as before, but results obtained by applying the software Silvaco. Injected current: $I_0 = 200mA$.

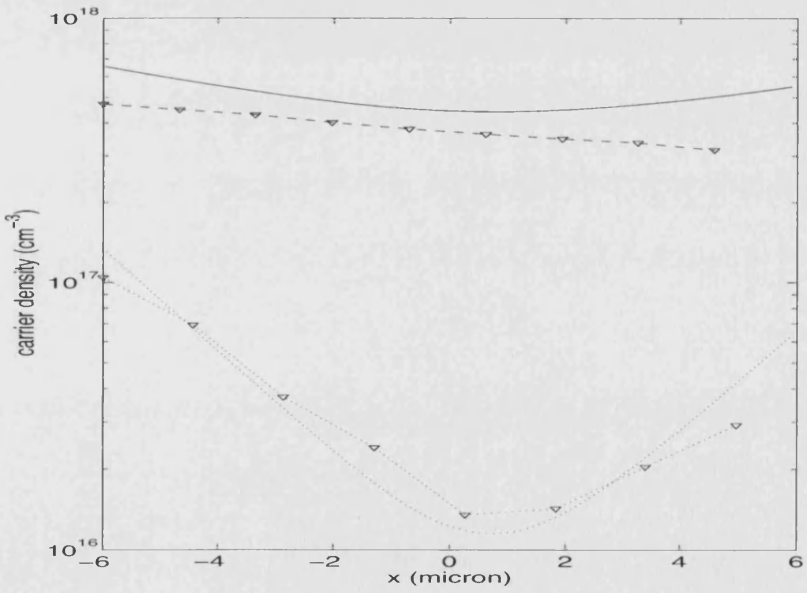


Figure 7.18: Carrier profile (in cm^{-3}) at the horizontal cross-section, $y = -1\mu m$, for two different values of carrier lifetime. Solid line: $\tau_{p0} = \tau_{n0} = 2 \cdot 10^{-7}s$; dotted line: $\tau_{p0} = \tau_{n0} = 2 \cdot 10^{-9}s$. The solid-diamond and dotted-diamond lines are for same lifetime constants as before, but results obtained by applying Silvaco. Injected current: $I_0 = 200mA$.

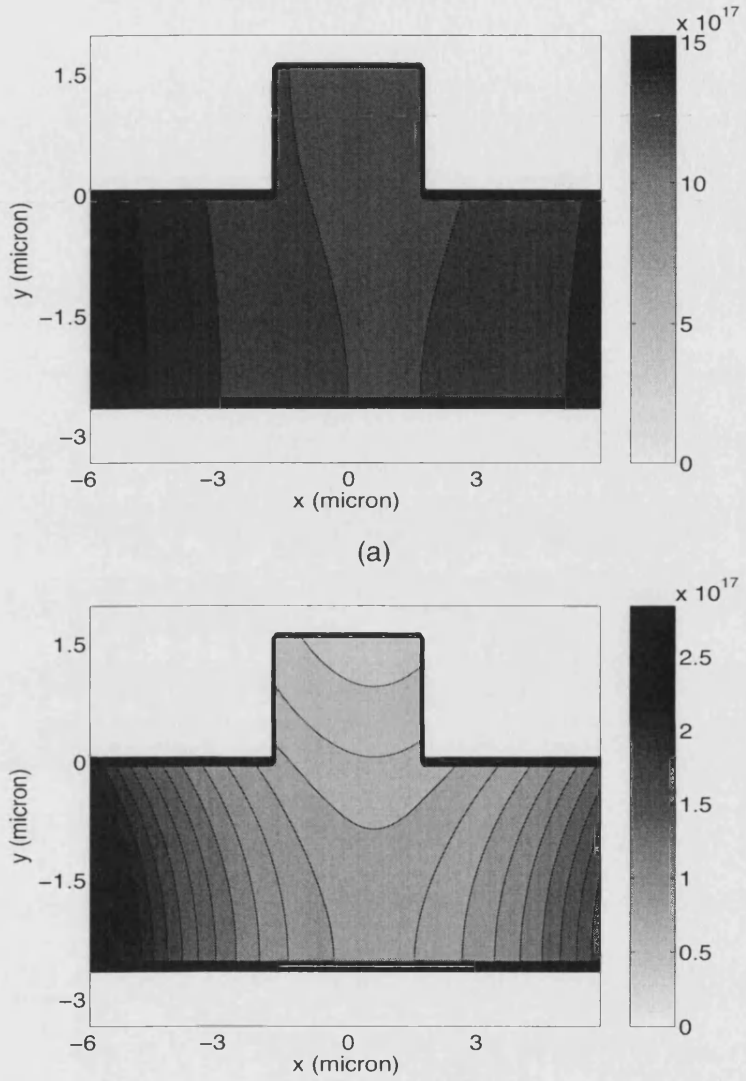


Figure 7.19: Effect of surface recombination on the carrier profile $p(x, y)$ (in cm^{-3}). (a) $S_r = 0$: surface recombinations are not included. The small vertical gradient is due only to the presence of the rib; (b) $S_r = 10^4 cm/s$. Carriers recombine faster in proximity of the top surface.

in Tab. [7.3].

In contrast to the design of the structure introduced in Fig. [7.1], the waveguide studied

dimensions:
$w = 3.35\mu m$, $l = 10\mu m$, $L = 500\mu m$
$d = 2.58\mu m$, $h = 1.59\mu m$
recombination coefficients:
$\tau_{p0} = 300ns$, $\tau_{n0} = 700ns$

Table 7.3: Dimensions and recombination coefficients for the waveguide analysed by Hewitt et al. in [5].

by Hewitt has highly doped regions which extend only a part away below the surface of the metal contacts, Fig. [7.20]. However, this difference can be included in the LIM by the inclusion of the term, $J_{pv}(y = -d)$, in eqn. (6.25), to account for the escape of carriers into the bottom Si region, Fig. [7.20].

Fig. [7.21a] compares the results obtained by the two methods, for the averaged refractive

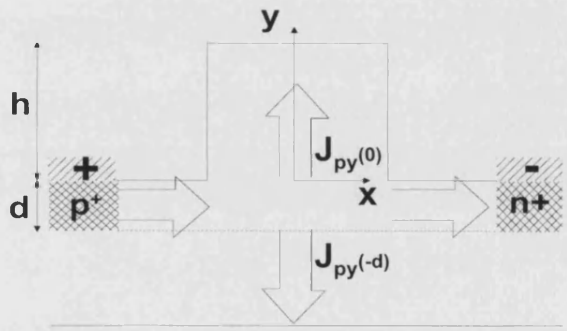
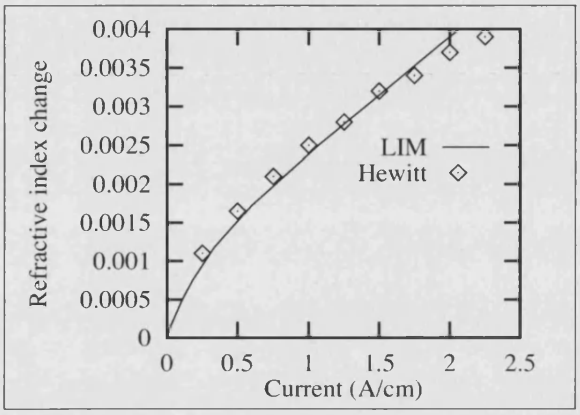


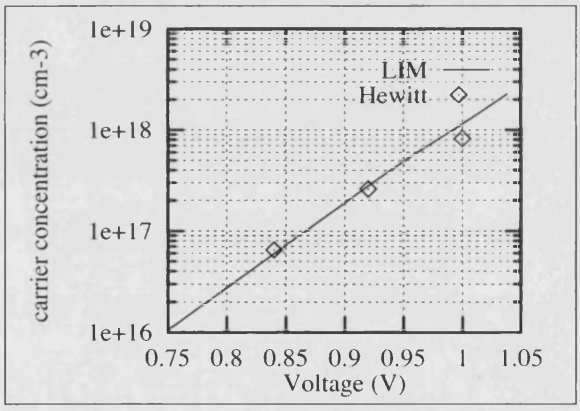
Figure 7.20: Structure of the waveguide analysed by Hewitt et al. in [5].

index change versus different values of injected current. The relation applied between carrier concentration and refractive index change is the same as used by Hewitt, already introduced by Soref in [6], eqn. (5.1).

The total voltage drop between the two contacts is determined by solving Poisson's eqn. (5.10) for the electric field corresponding to the carrier profile calculations. Results for carrier concentration versus voltage drop between the two contacts are compared in Fig. [7.21b]; the two methods provide results in very good agreement.



(a)



(b)

Figure 7.21: Comparison of results between LIM and Hewitt: (a) refractive index change for different values of injected current; (b) averaged carrier distribution versus voltage drop across the device.

Bibliography

[1] Grove A.S., "Physic and Technology of Semiconductor Devices", Wiley, 1967

[2] "Atlas - User's Manual. Device Simulation Software", SILVACO International, www.silvaco.com

[3] Sze S.M., Physics of Semiconductor Devices, John Wiley & Sons, NY, 1981

[4] "National Compound Semiconductor Roadmap", Dept. of the Navy, Science and Technology, www.ncsr.csci-va.com

[5] Hewitt P.D., Reed G.T., "Improving the Response of Optical Phase Modulators in SOI by the Computer Simulation", *J. Lightwave Technol.*, V.18, No.3, 2000, 443-450

[6] Soref R.A., Bennett B.R., "Electrooptical Effects in Silicon", *IEEE J. Quantum Electron.*, QE-23, No.1, 1987, 123-129

Chapter 8

Carrier controlled optical devices

As described in the Introduction, after injecting current, the consequent effect of the refractive index change on the field propagation must be determined. This prediction can be achieved either by solving again the wave equation, but applied to the altered refractive index profile, or by regarding the change of the refractive index as a perturbation to the initial (unperturbed) structure, Fig. [8.1].

The change of the refractive index reflects in the dielectric constant term $\epsilon(x, y, z)$ in the wave eqn. (2.10). After the perturbation the new refractive index profile may assume a non-uniform distribution across the device, and, as a consequence, the wave equation becomes cumbersome and difficult to be solved with the methods introduced in the first part of the thesis.

A valid and efficient alternative to this problem consists of considering the solutions of the original unperturbed waveguide and treating the changes as a perturbation.

After introducing the formalism of the perturbation theory, in this chapter some device configurations are analysed and given as examples of possible application of the models suggested in the thesis. The configurations illustrated are firstly solved in absence of any carrier injection, by utilising the electromagnetic models introduced in Chapter 3 and 4 for longitudinally uniform and non-uniform waveguides. The new refractive index distribution, after carrier injection, is then calculated by applying the Lumped Iterative Method. Finally the consequent changes in optical propagation are determined by applying the perturbation theory.

8.1 Perturbation analysis

The stationary perturbation theory determines the changes of the propagation characteristics of the modes when a small disturbance is applied, [1].

In the case of a longitudinally uniform waveguide the eigenvalue wave eqn. (2.15) for the

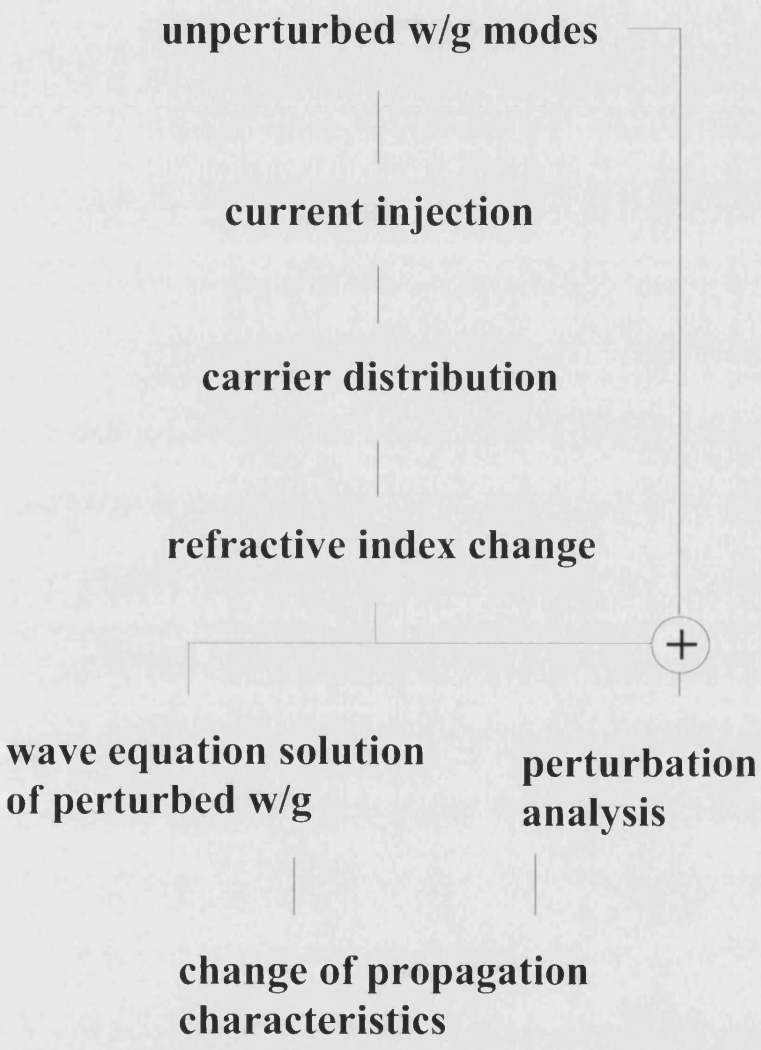


Figure 8.1: Flow chart describing different ways (in bold boxes) of analysing the effect of the refractive index change on the optical field propagation.

unperturbed waveguide can be written in a compact form as

$$U^0 E_n^0 = a_n^0 E_n^0 \tag{8.1}$$

with $U^0 = \partial_x^2 + \partial_y^2 + {k_0}^2 \epsilon_0(x,y)$ and $a_n^0 = (\beta_n^0)^2$; the superscript zero refers to unperturbed values, the subscript n refers to the n-th mode of the waveguide.
If the operator U^0 is altered by the perturbing operator u, the perturbed scalar wave equation can be rewritten in terms of the new operator $U = U^0 + u$, as

$$(U^0 + u)E_n = a_n E_n \tag{8.2}$$

where $E_n = E_n(x,y)$ and $a_n = \beta_n^2$ are the perturbed field and propagation constant, respectively.
Assuming that the perturbation term, u, is small it is possible to expand the perturbed propagation constant, a_n, as a power series in proximity of the unperturbed value, a_n^0.

$$a_n = a_n^0 + a_n^1 x + a_n^2 x^2 + a_n^3 x^3 + \ldots \tag{8.3}$$

The expansion coefficients a_n^1, a_n^2, a_n^3, etc., represent the first, second, third order of perturbation. In most cases it is sufficiently accurate to consider only perturbations of first order, i.e. evaluate only the coefficient a_n^1 in eqn. (8.3), [2]. The theory for the first order perturbation analysis is developed in Appendix G.
As a result, if the refractive index profile is changed from the unperturbed profile $\epsilon_0(x,y)$ to the new value $\epsilon(x,y) = \epsilon_0(x,y) + \Delta\epsilon(x,y)$, the effect of the perturbation $\Delta\epsilon(x,y)$ on the propagation constant, $\beta_n^2 = (\beta_n^0)^2 + \Delta\beta_n^2$, is

$$a_n^1 = \Delta\beta_n^2 = \frac{\int_S E_n^0 \Delta\epsilon(x,y) E_n^{0*} \mathrm{d}x\mathrm{d}y}{\int_S |E_n^0|^2 \mathrm{d}x\mathrm{d}y} \tag{8.4}$$

the integration is over the transverse spatial domain S.
In particular, if the perturbation of the dielectric inside the core region of a waveguide is constant, i.e. $\Delta\epsilon(x,y) = \bar{\Delta}$, then the perturbation on the β is proportional to the confinement factor of the field profile, Γ_{core}, inside the core region, i.e.,

$$\Delta\beta_n^2 = \bar{\Delta}\frac{\int_A E_n^0 E_n^{0*} \mathrm{d}x\mathrm{d}y}{\int_A |E_n^0|^2 \mathrm{d}x\mathrm{d}y} = \bar{\Delta}\Gamma_{core} \tag{8.5}$$

As indicated in Appendix G, in order to validate the perturbation theory, it is necessary that the modes of the unperturbed waveguide, E_n^0, are orthogonal. In the case of the two-dimensional rib waveguides solved by applying the Spectral Index Method it has been explicitly checked and verified numerically that the modes determined by the SIM are orthogonal.
Another important requirement to increase the accuracy of the perturbation analysis is the accuracy of the unperturbed field profile, $E_n^0(x,y)$. Because the effect of the perturbation

relies on the integration of the field profile, eqn. (8.4), the description of the profile must be numerically accurate. However, the SIM does satisfy this requirement, thus providing one more justification why the SIM is more favourable than other methods for determining 2-D optical field profiles.

8.2 A small library of SOI devices

The quasi-analytic models suggested in parts I and II of the thesis, combined with the perturbation analysis, are applied to a variety of device configurations to demonstrate how they can be used as an integrated Computer Assisted Design (CAD) tool to design novel devices.
The process of analysis can be summarised in the following three steps:

1. The field propagation of the unperturbed device (before injecting carriers, OFF state) is determined. The eigenmodes, $F_m^{off}(x,y)$, and propagation constants, β_m^{off}, of uniform parts of the device are determined by applying the Spectral Index Method (SIM) for longitudinal uniform waveguides, Chapter 3. Each mode propagates along the longitudinal z axis as

$$E_m^{off}(x,y,z) = F_m^{off}(x,y)\exp(-i\beta_m^{off}z) \tag{8.6}$$

   The total (unperturbed) field propagation along uniform sections of the device is given by

$$E^{off}(x,y,z) = \sum_m F_m^{off}(x,y)\exp(-i\beta_m^{off}z) \tag{8.7}$$

   The field propagation along non-uniform sections is determined by applying one of the SIM modified version for non-uniform waveguides, presented in Chapter 4.

2. The refractive index distribution, $\Delta\epsilon(x,y)$, consequent to the carrier injection is determined by applying the Lumped Iterative Method, introduced in Chapter 6.

3. The altered optical propagation (ON state) is predicted by applying the perturbation analysis. The new values of the propagation constants

$$\beta_m^{on} = \beta_m^{off} + \Delta\beta_m \tag{8.8}$$

   are calculated by using eqn. (8.4). These values are then used to determine the perturbed field distribution, as

$$E^{on}(x,y,z) = \sum_m F_m^{on}(x,y)\exp(-i\beta_m^{on}z) \tag{8.9}$$

The class of integrated devices which have been analysed and modelled, all based on Silicon-on-Insulator (SOI) technology, includes:

- Rib waveguide coupler
- Multimode Interferometer (MMI)
- Y-junction splitter

From the carrier injection analysis obtained in Chapter 7 it is possible to observe that, in the case of Silicon material, the carrier profiles are, in general, almost flat everywhere across the device. Although they are injected from the doped regions, carriers tend to diffuse largely (due to the large diffusion coefficients, D_p, D_n, of Silicon). These results impose a limit on the possibility of localising carriers inside Silicon.

So, for example, in the injection electrode configuration of Fig. [6.1b] there is evidence that a large amount of carriers diffuse also into the tail region B. Similarly, in the case of two ribs waveguides, e.g. Y-junction, coupler, Figs. [8.6], [8.2], if the metal contacts are deposited only at the sides of one rib, in order to alter the refractive index only across that rib, it will not be possible to confine the carriers locally. In fact a large amount of carriers will diffuse and spread also corresponding to the second rib, thus generating a wider spatial refractive index change.

8.2.1 Rib waveguide coupler

The configuration of a two ribs waveguide coupler of length L is depicted in Fig. [8.2]. As analysed in Section 3.5.1, the waveguide coupler supports two eigenmodes, namely the symmetric and anti-symmetric modes, Fig. [3.8], characterised by propagation constants, β_{sym} and β_{asym}, respectively. If the coupler is designed with a length L equivalent to the coupling length, L_c, i.e.,

$$L = L_c = \frac{\pi}{(\beta_{sym} - \beta_{asym})} \tag{8.10}$$

then, by launching the input field in to one branch, e.g. A, the optical power is entirely transferred to the other branch, B, at the output, Fig. [8.2]. If the length of the coupler is doubled, i.e. $L = 2L_c$ then the same input field is coupled to the parallel rib, B, first and then coupled back in to the original rib, A, at the output. Therefore, for a given device length L, by changing the coupling length it is possible to operate the device as a switch and select the desired output channel, either A or B.

The most effective way to change L_c is by altering one arm, e.g. A, of the coupler, injecting carriers from two metal contacts deposited at the sides of that branch (asymmetric contacts). However, in Silicon carriers will diffuse even toward the second branch, B, so that this configuration is practically not realisable. As alternative the metal contacts can be deposited at the sides of the two ribs (symmetric contacts), Fig. [8.2].

The injection of carriers produces a change on both the two propagation constants, $\beta_{sym}^{off} \rightarrow \beta_{sym}^{on}$ and $\beta_{asym}^{off} \rightarrow \beta_{asym}^{on}$, which, consequently, reflects on the coupling length, eqn. (8.10). Results for a coupler with dimensions as in Tab. [3.2], and a rib separation of $T = 2.48\mu m$ are shown in Tab. [8.1].

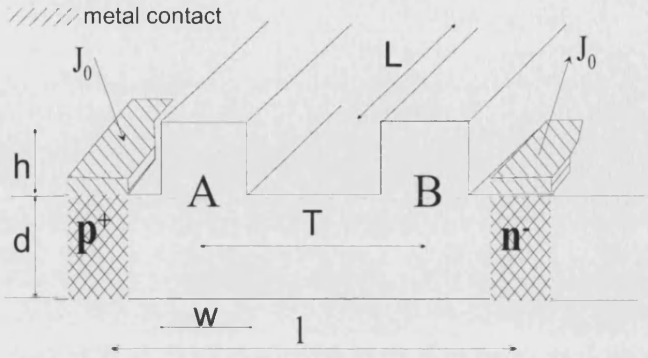


Figure 8.2: Scheme of a two ribs waveguide coupler.

J_0 (A/cm^2)	β_{sym}	β_{asym}	L_c (μm)
0	3.468309	3.467907	1924
200	3.468296	3.467886	1891
1000	3.468253	3.467824	1808
2000	3.468214	3.467731	1604
4000	3.468169	3.467512	1180

Table 8.1: Results for a coupler waveguide. J_0: injected current density; β_{sym} propagation constant for the symmetric mode; β_{asym} propagation constant for the anti-symmetric mode; L_c: coupling length.

8.2.2 Multimode Interferometer (MMI)

In the Multimode Interferometer (MMI) configuration, Fig. [8.3], the input beam is launched into a (multimode) waveguide section supporting more modes.

When carriers are not injected (OFF state) the field pattern excited at one side of the

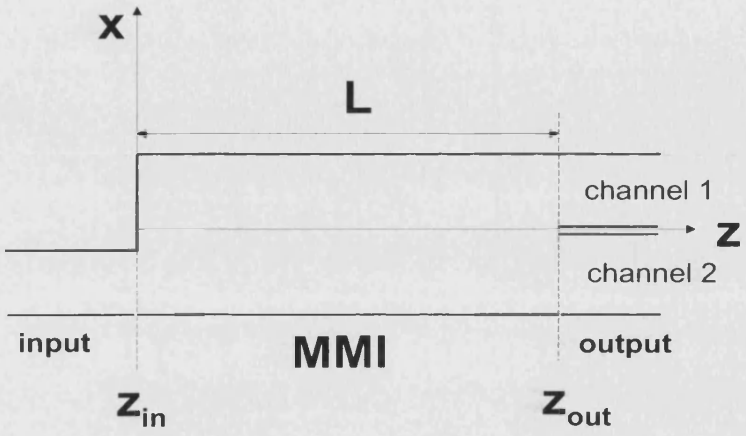


Figure 8.3: Scheme of a MMI.

device must emerge on either one of the two output channels. If the waveguide supports two guided modes, with propagation constants, β_1, β_2, a phase shift of $m\pi$ is necessary in order to produce a total field entirely confined at one output gate,

$$\Delta\beta^{off} \cdot L = (\beta_1^{off} - \beta_2^{off}) \cdot L = m\pi \tag{8.11}$$

L is the length of the multimode section, Fig. [8.3].

When the carriers are injected (ON state) the difference between the propagation constants of the modes changes. The consequence is a shift of the total field inside the multimode section, from one side to the other side. An additional π phase shift is then needed to steer the optical profile to the opposite output of the device:

$$\Delta\beta^{on} \cdot L = (\beta_1^{on} - \beta_2^{on}) \cdot L = (m+1)\pi \tag{8.12}$$

From eqns. (8.11), (8.12) it is determined the required length L of the device to achieve optical switching,

$$L = \frac{\pi}{\Delta\beta^{on} - \Delta\beta^{off}} \tag{8.13}$$

L is reduced if the injection of current induces a larger change on the differential of the two propagation constants for the ON state.

Fig. [8.4] shows the field propagation for the OFF state: at the input stage the field is confined in the left part of the multimode waveguide; at the output the optical field emerges

average injected carrier density: ON state: $p = 6.77 \cdot 10^{18}(cm^{-3})$
average refractive index change: ON state: $\Delta n = -1.5817 \cdot 10^{-2}$
propagation constants: OFF state: $\beta_1 = 14.068681$, $\beta_2 = 14.053735$, $\Delta\beta^{off} = 0.014946$ ON state: $\beta_1 = 14.004533$, $\beta_2 = 13.989523$, $\Delta\beta^{on} = 0.01501$
length $L = 49087\mu m$

Table 8.2: Results for a MMI device.

from the same side. Fig. [8.5] refers to the ON state: the input field is the same as in the OFF state, but at the output the optical field selects the opposite channel. The numerical

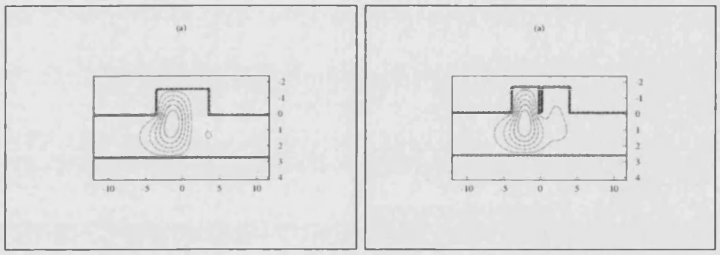


Figure 8.4: MMI: contour plot of the transverse field profile, (A.U.), for the OFF case, at the input and output stages, Fig. [8.3].

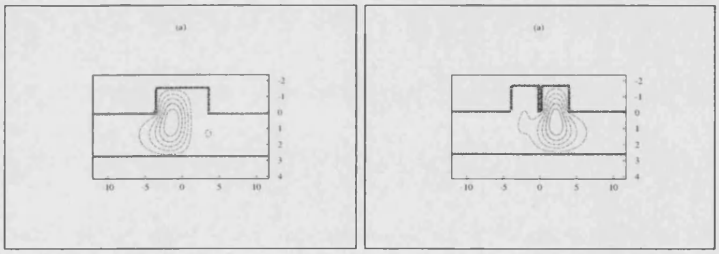


Figure 8.5: MMI: contour plot of the transverse field profile, (A.U.), for the ON case, at at the input and output stages, Fig. [8.3].

results are summarised in Tab. [8.2].

8.2.3 Y-junction splitter

The configuration of a Y-junction splitter is depicted in Fig. [8.6]. The light is coupled from an external source into the single rib section, (a), and propagated toward the adjacent (linear) tapering section, (b). At the apex, beginning of section (c), the optical field is split

equally in two parts and propagated through the two arms of the splitter. By lowering the refractive index of one arm it is possible to cause the optical suppression of that arm. As a consequence the optical field will be forced to propagate along the unaltered arm, hence causing optical switching.

An example of optical field propagation along a Y-junction with no arm suppression is

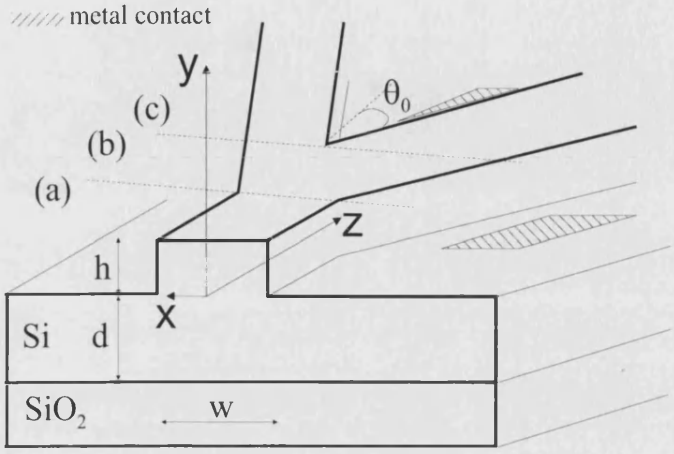


Figure 8.6: Scheme of a Y-junction splitter.

given in Figs. [4.11], [4.12], Section 4.9. As a comparison, the effect of an injected current (in the right arm of the splitter), and the consequent optical arm suppression is shown in Fig. [8.7], for the same device. The dimensions and parameters are the same as described in Tab. [3.2].

However, in the case of Silicon material, because carriers diffuse even toward the non-injected rib, it is necessary to deposit the metal contacts away from the apex, where the distance between the two arms is sufficiently large.

All the SOI devices presented in this section (coupler, MMI, Y-junction) can be solved by applying the software Silvaco. Since Silvaco is based on a Finite Element Method analysis it can support and solve the most complicated structure geometries. However, even for simple geometries, like these presented in this Section, the computational time required by Silvaco is of the order of some hours. Conversely, by combining the SIM for uniform/taper structures and the LIM it is possible to determine the electromagnetic field and the carrier distributions across the device with computational runtimes of few minutes (usually less than 15 minutes).

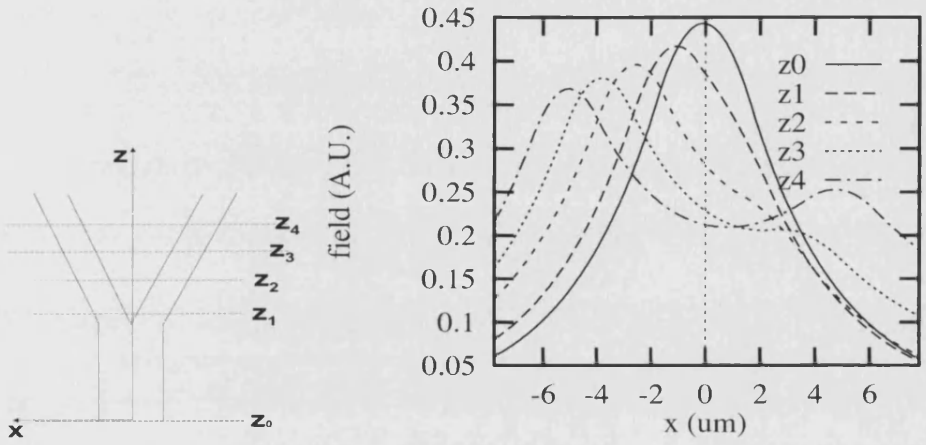


Figure 8.7: Optical propagation along a Y-junction after injecting carriers inside the right arm of the splitter. The lateral field profile (A.U.) is reported for different sections along the longitudinal z axis, $z_{0,1,2,3,4} = 0,\ 12,\ 30,\ 48,\ 66\ \mu m$. Angle $\theta_0 = 2^o$.

Bibliography

[1] Fermi E., "Notes on Quantum Mechanics", Phoenix Science Series, The University of Chicago Press, Chicago, 1961

[2] Kevorkian J., Cole J.D., "Perturbation methods in Applied Mathematics", Springer Verlag, NY, 1981

Chapter 9

Conclusion and Future work

9.1 Conclusion

The control of optical propagation inside planar dielectric waveguides has been investigated.
Quasi-analytic models capable of analysing complicated waveguide structures, whose solution is otherwise known only in terms of pure numerical techniques, have been suggested. The study has been conducted under two broad headings: the electromagnetic field propagation along deep ridge waveguides and the two-dimensional carrier distribution subsequent to current injection.
The optical field propagation along longitudinally non-uniform waveguides, such as taper and Y-junction configurations, has been predicted by using simple functional expansion techniques. The methods introduced constitute thus convenient tools to design and analyse the feasibility of novel structures.
An efficient and accurate method for determining two-dimensional carrier density profiles has been developed. The Lumped Iterative Method provides carrier and current distribution profiles inside two-dimensional structures with degrees of accuracy as those offered by more detailed commercial software packages.

The electromagnetic analysis focused mainly on the field propagation in deep ridge waveguides, i.e. characterised by a large refractive index step between the guiding region and the surrounding cladding region. That is the case e.g., of devices realised with the Silicon-on-Insulator (SOI) technology.
The well established quasi-analytic Spectral Index Method (SIM) has been utilised for the eigenmode propagation solution of longitudinally uniform waveguides. The use of the SIM provides results with an accuracy of the same extent of that obtained by using the pure numerical Finite Element Method (FEM). As a consequence the SIM can be used to over-

come the limits imposed by other more straightforward quasi-analytic methods (like the Effective Dielectric Constant), especially when applied to complex integrated structures. This is the case, for example, with a rib waveguide coupler, where the accuracy of the difference between the propagation constants of the modes is more important than their individual values in order to estimate the coupling length sufficiently accurately.

The SIM has also been extended and adapted to the analysis of longitudinally non-uniform (adiabatic) waveguides, like small angled tapers. The field propagation in taper devices can not be treated as an eigenmode problem, so the modified SIM has been applied for determining the longitudinal propagation of a specified input field launched at the start of the taper.
The field inside the taper has been represented with a judiciously chosen expression and the error introduced by such approximation minimised by applying a variational analysis. In this way it is possible to determine an analytic form for the slowly varying longitudinal field propagation.
Results have been compared with other numerical methods showing satisfactory agreement. The field profile spreads out laterally as the taper becomes larger, albeit it becomes effectively more confined inside the rib. The stronger influence of larger flare angles on the field propagation has also been investigated and illustrated.
Within devices with longitudinal variations the method has also been applied to Y-junction structures. In most published works, the propagation along Y-junction structures is analysed either by Local Mode Expansion technique or by using Finite Element/Beam Propagation methods. The approach presented in this thesis represents a unique alternative method to predict the longitudinal field decay due to the coupling to other guided or radiative modes.

Tapered rib waveguides have also been investigated by using metal horns eigenmodes as convenient set of basis functions for the functional expansion of the optical field. This approach constitutes a straightforward extension of the concepts utilised in the Spectral Index Method for longitudinally uniform waveguides to longitudinally non-uniform waveguides.

The optical propagation characteristics of a device can be altered by changing the refractive index of the semiconductor material. Amongst different possible ways of changing the refractive index the injection of carrier has been considered in the thesis. Therefore, the carrier distribution across the device, after injection of current, has been investigated.
As a first approximation the rib has been neglected and the planar structures treated as a one-dimensional p-i-n junction. However, in order to include the two-dimensional effects of the rib the Lumped Iterative Method (LIM) has been suggested as a simple, but accurate, method of solution.
The LIM reduces the two-dimensional problem, (x, y), to the separate solution of two

coupled one-dimensional problems in (x) and (y). The errors introduced by the approximations are minimised by iterating the process, and refining the solution at each successive cycle of iteration.
The carrier and current density (two-dimensional) profiles calculated by using the LIM have been compared with results obtained by the pure-numerical commercial software Silvaco, showing excellent agreement. Different recombination processes have been considered and represented in the model through the effective carrier lifetime constant. In addition, recombinations at surfaces have been introduced in the model as boundary conditions.
Results have been calculated for different values of injected current, recombination processes, doping densities, etc., matching with the results calculated from the software Silvaco. The presence of the rib has an effect on the carrier distributions as some of the carriers diffuse toward the rib, thus reducing the carrier concentration across the whole device. Moreover, the rib generates non-uniform carrier distributions, thus causing spatially localised refractive index changes.

The propagation of the optical field in the structure with the new (altered) refractive index profile can be determined by applying the Spectral Index Method (SIM). However, the SIM is not directly applicable to the case of waveguides with non-uniform refractive index profile. Therefore a Perturbation Analysis was introduced to predict the change of the optical field propagation following current injection. The waveguide is firstly solved for the unperturbed case and, then, the solutions used, together with the altered refractive index profile, to determine the new field solutions for the perturbed case.

Finally, examples of optical modulation in some waveguide structures are given, by applying all the different techniques described in the earlier chapters of the thesis. Different geometries have been simulated and presented, to prove the capability of the model as a flexible computed assisted design (CAD) tool.
In addition to providing a direct insight into the physics of the problem, the quasi-analytic methods introduced in this thesis offer also computational runtime advantages. In fact, most of the structures analysed were solved within few minutes, whereas, the same structures, required few hours computational runtime when solved by using the commercial software Silvaco (based on the pure numerical Finite Element Method).
Although all the numerical results presented in this thesis refers specifically to SOI devices, the models suggested are general enough to be applied to a wider variety of devices and materials, with minor modifications.

9.2 Future work

Further investigation of some topics presented in the thesis is envisaged, in order to refine the models introduced. Possible, most relevant, improvements are described in the following sections.

9.2.1 Field propagation in longitudinally non-uniform waveguides: functional expansion

The optical field inside the taper rib of a non-uniform waveguide can be, in general, represented by the functional expansion, eqns. (4.3), (4.4),

$$F(x,y,z) = \sum_{k=1}^{\infty} a_k(z) \cdot \cos\left[\frac{(2k-1)\pi}{w(z)}x\right] \cdot G_k(y) \cdot \exp(-ipz) \tag{9.1}$$

However, in the thesis only one term of the expansion (9.1) has been used. The simplified single term expression does not account for the coupling of the input field with other modes, i.e. guided and radiated, due to the taper effects. The variational analysis utilised, eqn. (C.17), must encompass such coupling, and describe the energy transfer from the propagating mode to other modes, so that the total energy is conserved.
A more accurate field representation can be obtained if more terms of expansion (9.1) are considered. In that case the coupling effect is automatically built in to the expression, and the variational analysis will be reduced to the same form of a uniform waveguide, eqn. (C.6). On the other hand, more coefficients $a_k(z)$, $G_k(y)$, in eqn. (9.1) need to be determined. The field expansion (9.1) produces a set of coupled linear differential equations, that, eventually, can be reduced to a matrix form,

$$\frac{\mathrm{d}a_k(z)}{\mathrm{d}z}G_k(y)+\frac{\mathrm{d}^2G_k(y)}{\mathrm{d}y^2}a_k(z)+\left[k_0^2\epsilon-p^2-\left(\frac{(2k-1)\pi}{w}\right)^2\right]a_k(z)G_k(y) = \sum_{l\neq k} b_l a_l(z) G_l(y) \qquad k=1,2,\ldots \tag{9.2}$$

In order to solve the set of equations (9.2) some approximations are needed, and the convenient choice of such assumptions is the main object to be targeted in future work.

9.2.2 Field propagation in longitudinally non-uniform waveguides: vectorial analysis

The field analysis in the thesis has been reduced, under valid approximations, to a scalar problem. However, the scalar solution can only describe five components of the electromagnetic field. If an accurate polarisation analysis is required then it is necessary to consider all the six vector components. As a consequence, all the three electric field components eqns. (2.8), together with the corresponding magnetic field equations must be solved.

The solution of six electromagnetic components is not straightforward and major changes may be needed to the models introduced in this thesis, although the approximation $\partial_z \epsilon = 0$ in eqn. (2.8) may still be retained in the case of adiabatic tapers or uniform waveguides.

9.2.3 Field propagation in longitudinally non-uniform waveguides: cylindrical coordinates

The quasi-modal analysis attempted in cylindrical coordinates by using metal horn eigenmodes is, perhaps, the most challenging and speculative work to be investigated further. The use of more terms to represent the field inside the rib, eqn. (4.23), is encouraged. But, more importantly, a detailed analysis of the use of the Hankel Transform, eqn. (4.31), is suggested. In fact, the consequent solution of the differential-integral eqn. (4.37) will certainly provide a deeper understanding of the coupling effects among different modes produced by the tapers.

9.2.4 Carrier dependent mobilities

The carrier transport equations (5.2), (5.3) have been solved without including explicitly the mobility dependence of the carrier concentration, i.e. $\mu_p = \mu_p(p)$, $\mu_n = \mu_n(n)$. Although in the iterative process the values of μ_p and μ_n have been readjusted according to the carrier concentrations obtained as result in the previous step of iteration, the mobilities have been considered constant within each cycle of iteration. Such artifice used in the model may not be accurate if the carrier densities are in the range where the mobility/ carrier concentration curve presents a large slope. That is the case, for example, of carrier concentrations in the range of $[10^{16} \div 10^{18}] cm^{-3}$, for Silicon, Fig. [7.2].

9.2.5 Refractive index change: electro-refraction, thermo-optical effect

Of all possible physical ways of changing the refractive index of a semiconductor emphasis has been given only to carrier injection. This particular restricted choice was dictated by the industrial requirement of analysing Silicon-on-Insulator (SOI) devices. However, as introduced in Chapter 5, the refractive index can also be altered by applying an external electric field or by exploiting the thermal properties of the material. Both these two cases can be accounted for by analysing the effects they produce on the quantum energy level distributions.

However, the temperature dependence analysis can also be conducted at a macroscopic level, by applying the same carrier transport equations (5.2), (5.3) and Poisson's eqn. (5.10) utilised in this thesis. The temperature effect is encompassed in the model by including the temperature dependence of the recombination rate, $R(T, p, n)$, the mobilities, $\mu_p(T)$, $\mu_n(T)$, and the diffusion coefficients, D_p, D_n.

The inclusion of electro-refraction and thermo-optical effects extends the applicability of the model to a wider variety of semiconductor materials, i.e. GaAs, Nitrides, etc.

9.2.6 Injection of carriers in new device configurations

The two-dimensional carrier distributions have been calculated for rib waveguides with both the metal contacts deposited on top of the core layer, Fig. [6.1]. This choice was particularly suitable for SOI devices, because of the bottom insulator layer. In general, the contacts may be deposited on both the top and bottom parts of the grown wafer (stripe lasers). This new configuration can be solved by extending the Lumped Iterative Method to the new structure, Fig. [9.1b].
Moreover the LIM can be easily extended also to other geometries, such as the waveguide coupler, Fig. [9.1a].

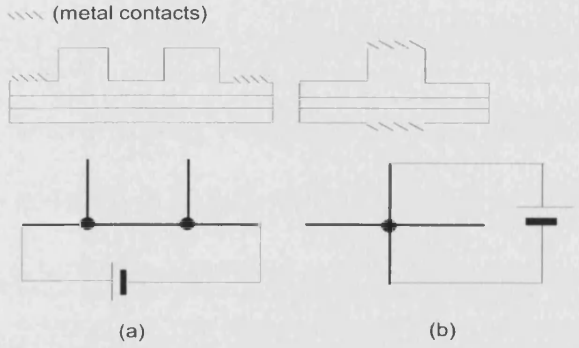


Figure 9.1: Lumped equivalent circuit for: (a) waveguide coupler; (b) stripe laser.

9.2.7 Dynamic analysis

In the present thesis the carrier transport eqns. (5.2), (5.3) have been solved in the steady state regime, i.e. without considering any time dependence. Since the main aim of the work was to analyse the final effects of the injected carriers on the refractive index, the time-independent analysis was acceptable. But, often, the time dependence is an important parameter in device modulation, especially in the case of fast modulation.
The dynamic analysis of a device is so cumbersome that may not be described exhaustively in the content of a thesis. However, if one particular dynamic aspect is targeted, i.e. transient time, stability analysis, etc., then it may be possible to make some specific approximations and obtain fairly accurate results with limited calculations.

9.2.8 Perturbation analysis: perturbation of the field profile

The perturbation theory has been applied specifically to determine the waveguide propagation constants, β, after altering the refractive index. Thus, it has been assumed that the main effect of the perturbation acts on the propagation constant, rather than the field profile, $E(x, y)$. In some cases, the perturbation may effect significantly even the field profile, and such changes can be determined by considering the coefficients c_{nm} in eqn. (G.13). Moreover, the perturbation analysis requires a complete set of (unperturbed) functions to use as a basis set for expanding the perturbed field profile. Such set is certainly available in the case of longitudinally uniform waveguide, but must be chosen carefully in the case of non-uniform waveguides. The application of the perturbation analysis to non-uniform waveguide is thus a topic which needs to be studied more in detail.

Part III

Appendix

Appendix A

Maxwell's equations and Vector Relations

A.1 Vector relations

f is a scalar variable

$\mathbf{A}$ is a vector

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla \nabla \cdot \mathbf{A} - \nabla^2 \mathbf{A} \tag{A.1}$$

$$\nabla \times (f\mathbf{A}) = \nabla f \times \mathbf{A} + f \nabla \times \mathbf{A} \tag{A.2}$$

$$\nabla \cdot (\nabla f) = \nabla^2 f \tag{A.3}$$

Cartesian Coordinates

unity vectors: $\hat{x}, \hat{y}, \hat{z}$

f is a scalar variable

$\mathbf{A} = (A_x, A_y, A_z)$ is a vector

$$\nabla f = \frac{\partial f}{\partial x}\hat{x} + \frac{\partial f}{\partial y}\hat{y} + \frac{\partial f}{\partial z}\hat{z} \tag{A.4}$$

$$\nabla^2 f = \frac{\partial^2 f}{\partial^2 x} + \frac{\partial^2 f}{\partial^2 y} + \frac{\partial^2 f}{\partial^2 z} \tag{A.5}$$

$$\nabla^2 \mathbf{A} = \nabla^2 A_x \hat{x} + \nabla^2 A_y \hat{y} + \nabla^2 A_z \hat{z} \tag{A.6}$$

Cylindrical Coordinates

unity vector: $\hat{r}, \hat{\theta}, \hat{y}$

f is a scalar variable

$\mathbf{A} = (A_r, A_\theta, A_y)$ is a vector

$$\nabla f = \frac{\partial f}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial f}{\partial \theta}\hat{\theta} + \frac{\partial f}{\partial y}\hat{y} \tag{A.7}$$

$$\nabla^2 f = \frac{\partial^2 f}{\partial^2 r} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial^2 \theta} + \frac{\partial^2 f}{\partial^2 y} \tag{A.8}$$

$$\nabla^2 \mathbf{A} = \left(\nabla^2 A_r - \frac{A_r}{r^2} - \frac{2}{r^2}\frac{\partial A_\theta}{\partial \theta}\right)\hat{r} + \left(\nabla^2 A_\theta - \frac{A_\theta}{r^2} + \frac{2}{r^2}\frac{\partial A_r}{\partial_\theta}\right)\hat{\theta} + \nabla^2 A_y \hat{y} \tag{A.9}$$

A.2 Maxwell's equations for dielectric materials

Charge free, $\rho = 0$, source free, $\mathbf{J} = 0$, material and harmonic time dependence are assumed.

Cartesian Coordinates

$$\partial_x \epsilon E_x + \partial_y \epsilon E_y + \partial_z \epsilon E_z = 0 \tag{A.10}$$

$$\partial_x H_x + \partial_y H_y + \partial_z H_z = 0 \tag{A.11}$$

$$\begin{aligned} -i\omega\mu_0 H_x &= \partial_y E_z - \partial_z E_y \\ -i\omega\mu_0 H_y &= \partial_z E_x - \partial_x E_z \\ -i\omega\mu_0 H_z &= \partial_x E_y - \partial_y E_x \end{aligned} \tag{A.12}$$

$$\begin{aligned} -i\omega\epsilon_0\epsilon E_x &= \partial_y H_z - \partial_z H_y \\ -i\omega\epsilon_0\epsilon E_y &= \partial_z H_x - \partial_x H_z \\ -i\omega\epsilon_0\epsilon E_z &= \partial_x H_y - \partial_y H_x \end{aligned} \tag{A.13}$$

Cylindrical Coordinates

$$\frac{1}{r}\partial_r r\epsilon E_r + \frac{1}{r}\partial_\theta \epsilon E_\theta + \partial_y \epsilon E_y = 0 \tag{A.14}$$

$$\frac{1}{r}\partial_r r H_r + \frac{1}{r}\partial_\theta H_\theta + \partial_y H_y = 0 \tag{A.15}$$

$$\begin{aligned} -i\omega\mu_0 H_r &= \tfrac{1}{r}\partial_\theta E_y - \partial_y E_\theta \\ -i\omega\mu_0 H_\theta &= \partial_y E_r - \partial_r E_y \\ -i\omega\mu_0 H_y &= \tfrac{1}{r}\partial_r r E_\theta - \tfrac{1}{r}\partial_\theta E_r \end{aligned} \tag{A.16}$$

$$\begin{aligned} -i\omega\epsilon_0\epsilon E_r &= \tfrac{1}{r}\partial_\theta H_y - \partial_y H_\theta \\ -i\omega\epsilon_0\epsilon E_y &= \partial_y H_r - \partial_r H_y \\ -i\omega\epsilon_0\epsilon E_z &= \tfrac{1}{r}\partial_r r H_\theta - \tfrac{1}{r}\partial_\theta H_r \end{aligned} \tag{A.17}$$

Appendix B

Evanescent boundary conditions

If the refractive index step between the core region of a rib waveguide (where a propagating guided field is bound) and the cladding region (where the guided field is evanescent) is large, the field decays rapidly into the cladding, and vanishes after a short distance away from the interface.

The decaying distance depends of the type of polarisation; that is, if the electric field is tangential, E_t, or perpendicular, E_n, to the interface.

With reference to Fig. [B.1] and coordinate system (t, n), the electric field inside the

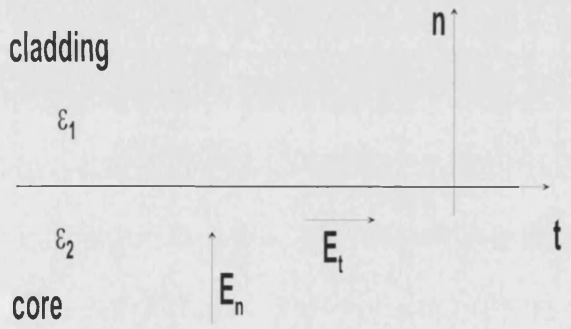


Figure B.1: Cladding/core interface.

cladding region satisfies Helmholtz's wave equation, eqn. (2.15), [1],

$$\frac{d^2E}{dn^2} + (k_0^2\epsilon_1 - \beta^2)E = 0 \tag{B.1}$$

with solution

$$E = E^{(1)} \exp(\sqrt{\beta^2 - k_0^2\epsilon_1}n) \tag{B.2}$$

$E^{(1)}$ is the value of the electric field at the interface $n = 0^+$, on the side of the cladding. It is then possible to establish two mathematical boundary conditions at $n = 0$ for the two

polarisations E_n, E_t

$$\frac{dE_n^{(1)}}{dn} = -\sqrt{\beta^2 - k_0^2\epsilon_1}E_n^{(1)} \tag{B.3}$$

$$\frac{dE_t^{(1)}}{dn} = -\sqrt{\beta^2 - k_0^2\epsilon_1}E_t^{(1)} \tag{B.4}$$

However, the field components, $E_n^{(1)}$, $E_t^{(1)}$ inside the cladding must be matched with the corresponding components inside the core, $E_n^{(2)}$, $E_t^{(2)}$, i.e.,

$$E_t{}^{(1)} = E_t{}^{(2)} \qquad \epsilon_1 E_n{}^{(1)} = \epsilon_2 E_n{}^{(2)} \tag{B.5}$$

and, similarly for their derivatives

$$\frac{dE_t{}^{(1)}}{dn} = \frac{dE_t{}^{(2)}}{dn} \qquad \frac{dE_n{}^{(1)}}{dn} = \frac{dE_n{}^{(2)}}{dn} \tag{B.6}$$

Substituting eqns. (B.5)-(B.6) into eqns. (B.3), (B.4) yields the evanescent boundary conditions

$$\frac{dE_n^{(2)}}{dn} = -\frac{E_n^{(2)}}{d_n} \tag{B.7}$$

$$\frac{dE_t^{(2)}}{dn} = -\frac{E_t^{(2)}}{d_t} \tag{B.8}$$

d_n and d_t represent the penetration depth of the field inside the cladding region before its amplitude becomes negligible

$$d_t = \frac{1}{\sqrt{\beta^2 - k_0^2\epsilon_1}}$$

$$d_n = \frac{\epsilon_1}{\epsilon_2}\frac{1}{\sqrt{\beta^2 - k_0^2\epsilon_1}} \tag{B.9}$$

In the case of a rib structure, with reference to Fig. [3.2], for TM_y polarisation the electric field, E_y, is tangential to the side walls of the rib and perpendicular to the top wall; for TE_y polarisation, the field E_x is tangential to the top wall and perpendicular to the vertical walls. Thus the dimensions of the effective rib in Fig. [3.2] are chosen either d_n or d_t according to the type of polarisation and the direction of the field with respect to the walls of the rib.

Appendix C

Variational method in the SIM

As introduced in Section 3.4 the assumptions made in the Spectral Index Method (SIM) do ensure the continuity of the field at the base of the rib, $y = 0$, Fig. [3.3], but they cause a mismatch of the field derivative, [1]. A variational technique is then introduced in order to minimise the error induced on the propagation constant β by the approximate method.

C.1 Variational analysis for uniform waveguides

In the case of a longitudinally uniform rib waveguide the exact field solution satisfies Helmholtz's equation, eqn. (2.15),

$$\frac{\partial^2 E(x,y)}{\partial x^2} + \frac{\partial^2 E(x,y)}{\partial y^2} + [{k_0}^2 \epsilon(x,y) - \beta^2] E(x,y) = 0 \tag{C.1}$$

Eqn. (C.1) can also be written in the variational form for the propagation constant β, [1], [2], [3],

$$\beta^2 = \frac{\iint_{A,B} \left[k_0^2 \epsilon |E|^2 - \left|\frac{\partial E}{\partial x}\right|^2 - \left|\frac{\partial E}{\partial y}\right|^2 \right] \mathrm{d}x\mathrm{d}y}{\iint_{A,B} |E|^2 \mathrm{d}x\mathrm{d}y} \tag{C.2}$$

The integrals are over the entire transverse section, i.e. the union of the two regions A (top) and B (bottom), Fig. [3.3].

If the field representation $E(x,y)$ is not an exact solution (as in the case of the SIM) the variational form which makes β stationary assumes a different form than eqn. (C.2). Multiplying Helmholtz's eqn. (C.1) by $E^*(x,y)$ (complex conjugate) and integrating over each region A and B, it is possible to obtain a new variational form for the propagation constant β. More precisely, it is

$$\beta^2 = \frac{\iint_A \left[E_A^* \frac{\partial^2 E_A}{\partial x^2} + E_A^* \frac{\partial^2 E_A}{\partial y^2} + k_0^2 \epsilon_2 E_A E_A^* \right] \mathrm{d}x\mathrm{d}y}{\iint_A E_A E_A^* \mathrm{d}x\mathrm{d}y} \tag{C.3}$$

in region A, and

$$\beta^2 = \frac{\iint_B \left[E_B^* \frac{\partial^2 E_B}{\partial x^2} + E_B^* \frac{\partial^2 E_B}{\partial y^2} + k_0^2 \epsilon_2 E_B E_B^*\right] \mathrm{d}x\mathrm{d}y}{\iint_B E_B E_B^* \mathrm{d}x\mathrm{d}y} \tag{C.4}$$

in region B.

Integrating eqn. (C.3) by parts, and using the divergence theorem [1], after some algebra, yields

$$\beta^2 = \frac{\iint_A \left[k_0^2 \epsilon_2 |E_A|^2 - \left|\frac{\partial E_A}{\partial x}\right|^2 - \left|\frac{\partial E_A}{\partial y}\right|^2\right] \mathrm{d}x\mathrm{d}y - \int_{C_A} E_A^* \frac{\partial E_A}{\partial x} \mathrm{d}l - \int_{C_A} E_A^* \frac{\partial E_A}{\partial y} \mathrm{d}l}{\iint_A E_A E_A^* \mathrm{d}x\mathrm{d}y} \tag{C.5}$$

The second and third line integrals at the numerator of eqn. (C.5) are calculated around the boundary of region A, as indicated by the closed (dotted) curve C_A in Fig. [C.1]. Because the field is zero, $E = 0$, along the rib, these two integrals reduce to the integration along the x axis (the only part of the curve C_A where the field is non-zero). A similar form of eqn. (C.5) is obtained by solving the integral eqn. (C.4) in region B; in that case, since the field vanishes at $y = -\infty$ the line integrals around C_B (dashed line in Fig. [C.1]) reduce to the integration along the x axis.

Adding (C.3)-(C.4) and making use of eqn. (C.2) yields

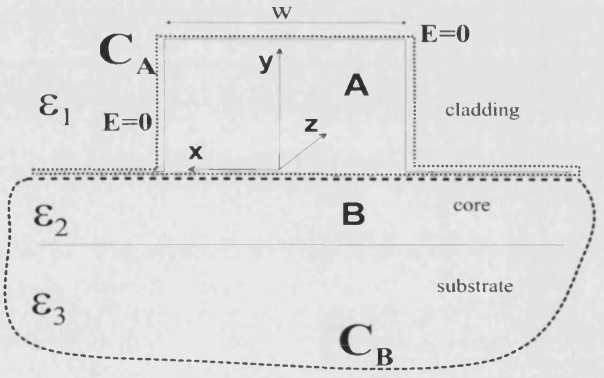


Figure C.1: Regions of integration. C_A (dotted line): boundary of region A; C_B (dashed line): boundary of region B.

$$\int_{-\infty}^{+\infty} \left[E_A^* \frac{\partial E_A(x, y = 0^+)}{\partial y} - E_B^* \frac{\partial E_B(x, y = 0^-)}{\partial y}\right] \mathrm{d}x = 0 \tag{C.6}$$

It is useful to note that if the field expression $E(x, y)$ is an exact solution, i.e. the continuity of the field derivative at the base of the rib can be ensured, than eqn. (C.6) will be an identity and eqns. (C.3), (C.4) reduce to the exact form (C.2).

[1] Divergence theorem, [4]: Let S be a region in space whose boundary is a piecewise smooth closed curve l, and $\mathbf{F}(x, y)$ a continuous function with continuous first partial derivatives. Then $\int \int_S \frac{\mathrm{d}\mathbf{F}}{\mathrm{d}x} + \frac{\mathrm{d}\mathbf{F}}{\mathrm{d}y} \mathrm{d}x\mathrm{d}y = -\int_l \mathbf{F} \cdot \mathbf{n} \mathrm{d}l$, where $\mathbf{n}$ is the outer unit normal vector of l

Although in the SIM the two derivatives in eqn. (C.6) are expressed in different domains, namely $\partial_y E_A(x, y = 0^+)$ in the spatial domain and $\partial_y E_B(x, y = 0^-)$ in the spectral domain, they can be linked by using Parseval's identity, [5]. Therefore, considering that the field outside the effective rib is zero, i.e. $E_A = 0$ for $|x| > w/2$, Fig. [C.1], eqn. (C.6) becomes

$$\int_{-w/2}^{w/2} E_A^* \frac{\partial E_A(x, y = 0^+)}{\partial y} \mathrm{d}x = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathcal{E}_B^* \frac{\partial \mathcal{E}_B(x, y = 0^-)}{\partial y} \mathrm{d}s \tag{C.7}$$

$\mathcal{E}_B$ is the Fourier Transform of $E_B(x, y)$ along the x axis, eqn. (3.12).

C.2 Dispersion relation for uniform waveguides

Eqn. (C.7) is the dispersion relation of a two-dimensional rib waveguide, whose roots are the sought propagation constants β.

Evaluation of the left hand side (LHS) of eqn. (C.7)

By imposing the field to be zero at $y = h$, Fig. [3.3], $G(y)$ in eqn. (3.9) becomes

$$G(y) = A_1 \frac{\sin k_{2A}(h - y)}{\sin k_{2A} h} \tag{C.8}$$

with k_{2A} defined as in (3.10). A_1 is an arbitrary constant which can be chosen to be unity for convenience, $A_1 = 1$, since the field profile amplitude is calculated in arbitrary units (A.U.). Using eqn. (C.8) and the expression for E_A in eqn. (3.7) the LHS of eqn. (C.7) becomes

$$\int_{-w/2}^{w/2} E_A(x, y = 0) \frac{\partial E_A^*(x, y = 0)}{\partial y} \mathrm{d}x = \frac{\pi}{2 s_A} k_{2A} \cot k_{2A} h \tag{C.9}$$

Evaluation of the right hand side (RHS) of eqn. (C.7)

The interface conditions (3.17) are applied at the common interface $y = -d$, Fig. [3.4], to the expressions of $\mathcal{E}_B(s, y)$ in eqn. (3.15):

- Continuity of the field, (3.17a):

$$-B_1 \sin(k_{2B} d) + B_2 \cos(k_{2B} d) = b C_1 \exp(-k_{3B} d) \tag{C.10}$$

  $b = 1$ or $b = \epsilon_3/\epsilon_2$ for TE_y, TM_y polarisation, respectively.

- Continuity of the field derivative, (3.17b):

$$k_{2B}[B_1 \cos(k_{2B} d) + B_2 \sin(k_{2B} d)] = k_{3B} C_1 \exp(-k_{3B} d) \tag{C.11}$$

The continuity of the field is also ensured at the base of the rib, $y = 0$, by making use of eqn. (C.8) in eqn. (3.20). This yields

$$B_2 = \int_{-w/2}^{w/2} \cos(s_A x) \exp(isx) \mathrm{d}x = \frac{2 s_A \cos(sw/2)}{s_A^2 - s^2} \tag{C.12}$$

Eqns. (C.10), (C.11) and (C.12) define uniquely the constants B_1, B_2, C_1:

$$\begin{aligned} B_1 &= \frac{b k_{2B} \sin(k_{2B} d) - k_{3B} \cos(k_{2B} d)}{b k_{2B} \cos(k_{2B} d) + k_{3B} \sin(k_{2B} d)} B_2 \\ B_2 &= \frac{2 s_A \cos(sw/2)}{s_A^2 - s^2} \\ C_1 &= b(B_1 \sin(k_{2B} d) + B_2 \cos(k_{2B} d)) \exp(k_{3B} d) \end{aligned} \tag{C.13}$$

and the consequent expression for $\mathcal{E}_B(s, y)$ as given in eqn. (3.15). Therefore, the RHS of eqn. (C.7) is

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathcal{E}_B(x, y = 0) \frac{\partial \mathcal{E}_B^*(x, y = 0)}{\partial y} \mathrm{d}s = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Gamma(s) \frac{4 s_A^2 \cos^2(sw/2)}{(s_A^2 - s^2)^2} \mathrm{d}s \tag{C.14}$$

with $\Gamma(s) = k_{2B} \frac{b k_{2B} \sin(k_{2B} d) - k_{3B} \cos(k_{2B} d)}{b k_{2B} \cos(k_{2B} d) + k_{3B} \sin(k_{2B} d)}$

The combination of eqns. (C.9) and (C.14) in eqn. (C.7) yields the dispersion relation for the propagation constant β,

$$k_{2B} \cot(k_{2B} h) = \frac{4 s_A^3}{\pi^2} \int_{-\infty}^{+\infty} \Gamma(s) \frac{\cos^2(sw/2)}{(s_A^2 - s^2)^2} \mathrm{d}s \tag{C.15}$$

C.3 Variational analysis for non-uniform waveguides

The slowly varying term of the field profile, $F(x, y, z)$ in eqn. (4.2), satisfies the paraxial wave equation

$$\left[\frac{\partial^2}{\partial_x^2} + \frac{\partial^2}{\partial y^2} - 2ip\frac{\partial}{\partial z} + k_0^2 \epsilon(x, y) - p^2\right] F(x, y, z) = 0 \tag{C.16}$$

Multiplying eqn. (C.16) by F^* (complex conjugate) and integrating yields

$$p^2 = \frac{\iiint_V \left[F^* \frac{\partial^2 F}{\partial x^2} + F^* \frac{\partial^2 F}{\partial y^2} + k_0^2 \epsilon F F^*\right] \mathrm{d}x\mathrm{d}y}{\iiint_V F F^* \mathrm{d}x\mathrm{d}y} - \frac{\iiint_V 2ipF^* \frac{\partial F}{\partial z} \mathrm{d}x\mathrm{d}y\mathrm{d}z}{\iiint_V F F^* \mathrm{d}x\mathrm{d}y\mathrm{d}z} \tag{C.17}$$

The integrals are over all of the volume space V occupied by the tapered rib waveguide. The first term of eqn. (C.17) can be treated equivalently to eqns. (C.3)-(C.4) for the uniform case. Integrating by parts and after some algebra the variational expression for a taper device is given by

$$\int_{-\infty}^{+\infty} \int_{0}^{+\infty} \left[F^* \frac{\partial F(x, y = 0^+)}{\partial y} - F^* \frac{\partial F(x, y = 0^-)}{\partial y}\right] \mathrm{d}x\mathrm{d}z + \frac{\iiint_V 2ipF^* \frac{\partial F}{\partial z} \mathrm{d}x\mathrm{d}y\mathrm{d}z}{\iiint_V F F^* \mathrm{d}x\mathrm{d}y\mathrm{d}z} = 0 \tag{C.18}$$

Appendix D

Slowly varying term for tapered geometry

The field profile inside the taper rib in Fig. [4.1] has been represented as, eqn (4.5),

$$F_A(x,y,z) = a(z) \cdot \cos\left[\frac{\pi}{w(z)}x\right] \cdot G(y) \tag{D.1}$$

Inserting eqn. (D.1) into the paraxial wave eqn. (2.19), multiplying by $\cos\left(\frac{\pi}{w}x\right)$ and integrating over x yields

$$\frac{\mathrm{d}^2G(y)}{\mathrm{d}y^2}a(z) + (k_0^2\epsilon - s_A^2 - p^2)G(y)a(z) = 2ip\left[\frac{\mathrm{d}a(z)}{\mathrm{d}z} + \frac{\tan\theta_0}{w(z)}a(z)\right]G(y) \tag{D.2}$$

with $s_A = \frac{\pi}{w_0}$.

In eqn. (D.2) it is possible to separate the two variables y and z and obtain two interdependent equations for $G(y)$ and $a(z)$ linked through the common complex parameter $q = q_R - iq_I$

$$\frac{\mathrm{d}^2G(y)}{\mathrm{d}y^2} + (k_0^2\epsilon - s_A^2 - p^2)G(y) = qG(y) \tag{D.3}$$

$$\frac{\mathrm{d}a(z)}{\mathrm{d}z} + \frac{\tan\theta_0}{w(z)}a(z) = -i\frac{q}{p}a(z) \tag{D.4}$$

The first equation (D.3) provides results similar to the uniform case, eqn. (3.9),

$$G(y) = A_1\sin(k_{2A}y) + A_2\cos(k_{2A}y) \tag{D.5}$$

with

$$k_{2A}^2 = k_0^2\epsilon_1 - s_A^2 - p^2 - q \tag{D.6}$$

and constants ratio A_1/A_2 to be determined by applying the boundary condition $E_A(x, y = h, z) = 0$.

The second equation (D.4) provides an analytic form for the slowly varying coefficient $a(z)$, i.e.,

$$a(z) = \frac{1}{\sqrt{w}}\exp\left(-i\frac{q}{p}z\right) \tag{D.7}$$

D.1 Slowly varying term for the Y-junction configuration

The field profile inside the rib arms of the Y-Junction is expressed as, eqns. (4.18), (4.19), Fig. [4.10],

$$E_A(x,y,z) = a(z)\cdot G(y)\cdot \begin{cases} A_1 \cos\left[\frac{(2k-1)\pi}{w}(x+T/2)\right] & (-w-T)/2 < x < (w-T)/2 \\ A_1 \cos\left[\frac{(2k-1)\pi}{w}(x-T/2)\right] & (T-w)/2 < x < (w+T)/2 \\ 0 & elsewhere \end{cases} \tag{D.8}$$

for symmetric fields, and

$$E_A(x,y,z) = a(z)\cdot G(y)\cdot \begin{cases} A_1 \cos\left[\frac{(2k-1)\pi}{w}(x+T/2)\right] & (-w-T)/2 < x < (w-T)/2 \\ A_1 \sin\left[\frac{2k\pi}{w}(x-T/2)\right] & (T-w)/2 < x < (w+T)/2 \\ 0 & elsewhere \end{cases} \tag{D.9}$$

for anti-symmetric fields. Inserting eqns. (D.8) (D.9) into the paraxial wave eqn. (2.19), multiplying by $\cos\left(\frac{\pi}{w}x\right)$ and integrating over x yields

$$\frac{\mathrm{d}^2G(y)}{\mathrm{d}y^2}a(z) + (k_0^2\epsilon - s_A^2 - p^2)G(y)a(z) = 2ip\frac{\mathrm{d}a(z)}{\mathrm{d}z}G(y) \tag{D.10}$$

with $s_A = \frac{\pi}{w}$.

The separation of the x and z variables in eqn. (D.10) yields explicitly the expressions for $G(y)$ and $a(z)$. Namely, the form of $G(y)$ is the same as in eqns. (D.5), (D.6), while the slowly varying coefficient $a(z)$ is obtained as

$$a(z) = \exp\left(-i\frac{q}{p}z\right) \tag{D.11}$$

Appendix E

Carrier concentration solution in p-i-n junctions

E.1 Intrinsic region

As introduced in Section. 6.2 in the case of a one-dimensional p-i-n junction the current conservation eqns. (5.2), (5.3) are reduced to

$$\frac{\mathrm{d}J_p(x)}{dx} = -eR_p = -e\frac{p(x)}{\tau_p} \tag{E.1}$$

$$\frac{\mathrm{d}J_n(x)}{dx} = eR_n = e\frac{n(x)}{\tau_n} \tag{E.2}$$

Recombination processes, R_p, R_n, have been represented through the effective lifetimes τ_p, τ_n, as defined in eqn. (5.12).

Moreover, inside the intrinsic region both the *high-level* injection and the *neutrality* conditions are valid. Under these assumptions it is then possible to solve the carrier transport eqns. (E.1), (E.2). Specifically, multiplying (E.1) and (E.2) by μ_n and μ_p, respectively, and adding yields the ambipolar equation,

$$D^*\frac{\mathrm{d}^2p(x)}{dx^2} + \mu^* E_x p(x) = \frac{p(x)}{\tau} \tag{E.3}$$

with

$$\frac{1}{\tau} = \frac{1}{\tau_p} + \frac{1}{\tau_n} \tag{E.4}$$

$$D^* = \frac{(n+p)D_nD_p}{nD_n + pD_p} \tag{E.5}$$

$$\mu^* = \frac{(n_0 - p_0)\mu_n\mu_p}{n\mu_n + p\mu_p} \tag{E.6}$$

the ambipolar effective lifetime, the ambipolar diffusion length and the ambipolar mobility, respectively.

Inside the intrinsic region it is $p_0 = n_0$, thus $\mu^* = 0$. Moreover, under the assumption of *high-level* injection it is, [7],

$$D^* = \frac{2\mu_n\mu_p}{\mu_n + \mu_p} \tag{E.7}$$

Since D^* is typically almost constant, analytic solutions for eqn. (E.3) are easily obtained as

$$p(x) = C_1 \cosh\frac{x}{L_i} + C_2 \sinh\frac{x}{L_i} \tag{E.8}$$

with $L_i = \sqrt{D^*\tau}$ the ambipolar diffusion length. The arbitrary constants C_1 and C_2 are determined by applying the boundary conditions at the two ends of the intrinsic region

$$J_n(x = -l/2) = e\mu_n n(-l/2)E(-l/2) + eD_n\frac{\mathrm{d}n}{\mathrm{d}x} = 0 \tag{E.9}$$

$$J_p(x = l/2) = e\mu_p p(l/2)E(l/2) - eD_p\frac{\mathrm{d}p}{\mathrm{d}x} = 0 \tag{E.10}$$

The only current flowing through the p-i junction, $x = -l/2$, Fig. [6.2], is the hole current, thus

$$J_p(-l/2) = J_0 = e\mu_p p(-l/2)E(-l/2) - eD_p\frac{\mathrm{d}p(x = -l/2)}{\mathrm{d}x} \tag{E.11}$$

where J_0 is the injected current density.

The electric field may then be determined from (E.9) as

$$E(-l/2) = \frac{-D_n}{\mu_n p(-l/2)}\frac{\mathrm{d}p}{\mathrm{d}x}(x = -l/2) \tag{E.12}$$

Inserting eqn. (E.12) into eqn. (E.11) it is possible to express the boundary condition at $x = -l/2$ explicitly in terms of the carrier density $p(x)$:

$$\frac{\mathrm{d}p}{\mathrm{d}x}(x = -l/2) = \frac{-J_0}{2e\mu_p V_t} \tag{E.13}$$

V_t is the thermal voltage [1].

In a similar manner it is possible to express the boundary condition at the i-n junction, $x = l/2$, Fig. [6.2], after determining the electric field from eqn. (E.10),

$$E(l/2) = \frac{D_p}{\mu_p p(l/2)}\frac{\mathrm{d}p(x = l/2)}{\mathrm{d}x} \tag{E.14}$$

The only current flowing at $x = l/2$ is the electron current, i.e.,

$$J_n(l/2) = J_0 = e\mu_n p(l/2)E(l/2) - eD_n\frac{\mathrm{d}p(x = l/2)}{\mathrm{d}x} \tag{E.15}$$

hence, the second boundary condition for $p(x)$ is

$$\frac{\mathrm{d}p}{\mathrm{d}x}(x = l/2) = \frac{J_0}{2e\mu_n V_t} \tag{E.16}$$

[1] $V_t = 0.026V$ at 300K

Applying the conditions (E.13),(E.16) to the expression (E.8) of $p(x)$ the two arbitrary constants C_1 and C_2 are uniquely determined as

$$C_1 = \frac{J_0}{2eV_tD_p}\frac{\mu_n+\mu_p}{\mu_n}\frac{L_i}{\sinh\frac{l/2}{L_i}} \tag{E.17}$$

$$C_2 = -\frac{J_0}{2eV_tD_p}\frac{\mu_n-\mu_p}{\mu_n}\frac{L_i}{\cosh\frac{l/2}{L_i}} \tag{E.18}$$

E.2 Doped regions

Across the junctions, Fig. [6.2], the *charge neutrality* condition is not valid: these regions have a net space-charge, yielding the generation of an electric field. Within the space-charge regions the electric field and the carriers are more strongly coupled. Particular attention is then necessary when developing numerical techniques for determining the carrier distribution at the junctions.

The physical boundaries of the doped regions are $x = -\infty$, $x = -l/2$, and $x = l/2$, $x = +\infty$ for the p and n side, respectively.

However, as mentioned in Sect. 6.2.1, it is not possible to use boundary conditions at infinity. Therefore an alternative approach is proposed which transforms the boundary value problem to an initial value problem. This implies that all the boundary conditions are assigned at the beginning of the doped regions, i.e. $x = -l/2$ and $x = l/2$ for the p and n regions, respectively. The values of $p(x)$ and $n(x)$ at the edge of the intrinsic region are obtained by using the analytic solution described in (E.8).

$$p(-l/2^-) = p(l/2^+) \qquad n(l/2^+) = n(l/2^-) \tag{E.19}$$

$$\frac{\mathrm{d}p(-l/2^-)}{\mathrm{d}x} = \frac{\mathrm{d}p(-l/2^+)}{\mathrm{d}x} \qquad \frac{\mathrm{d}n(l/2^+)}{\mathrm{d}x} = \frac{\mathrm{d}n(l/2^-)}{\mathrm{d}x} \tag{E.20}$$

Each region, p, n, is discretised into several small segments. Within each segment eqns. (6.1), (6.2) are linearised and solved analytically,

$$D_p\frac{\mathrm{d}^2p(x)}{\mathrm{d}x^2} - \mu_pE(x)\frac{\mathrm{d}p(x)}{\mathrm{d}x} - p(x)\left[\mu_p\frac{\mathrm{d}E(x)}{\mathrm{d}x} - \frac{1}{\tau_p}\right] = 0 \tag{E.21}$$

$$D_n\frac{\mathrm{d}^2n(x)}{\mathrm{d}x^2} + \mu_nE(x)\frac{\mathrm{d}n(x)}{\mathrm{d}x} + n(x)\left[\mu_n\frac{\mathrm{d}E(x)}{\mathrm{d}x} - \frac{1}{\tau_n}\right] = 0 \tag{E.22}$$

The algorithm is described as follows, for the p doped region, Fig. [E.1]:

1. The boundary conditions are assigned by evaluating the carrier profiles at the edges of the intrinsic region, $x = -l/2$, Fig. [6.2], using the results obtained in the intrinsic region.

2. The majority carrier profile, $p(x)$, inside one segment is determined by solving eqn. (E.21). The electric field, E, and its derivative, $\mathrm{d}E(x)/\mathrm{d}x$ in eqn. (E.21) are assumed to be constant within the segment (the constant value is taken to be that at the beginning of the segment). Hence the differential eqn. (E.21) reduces to a linear second order ordinary differential equation whose solution for the mth segment is in the same form as (E.8).

3. Similarly, the profile of the minority carrier, $n(x)$, could be determined by solving eqn. (E.22). Because at the junction minority carrier densities change more significantly than majority carriers the slope of $n(x)$ is expected to be large. An attempt to determine $n(x)$ in this way has been tried, but the algorithm resulted unstable, since the profile of $n(x)$ did not converge to the equilibrium value n_0 away from the junction. Therefore, the electron concentration has been solved by applying the total current conservation equation

$$\frac{dJ_{tot}}{dx} = \frac{d}{dx}(J_n + J_p) = 0 \tag{E.23}$$

Expressing the minority current J_n in terms of the carrier profile $n(x)$ as given in eqn. (5.6) it is

$$eD_n \frac{dn(x)}{dx} + e\mu_n E(x) n(x) = J_0 - J_p(x) \tag{E.24}$$

where J_p is estimated by using the values of $p(x)$ obtained from step (2).

4. The values of the carrier profiles, $p(x)$, $n(x)$, obtained from steps (2) and (3) are substituted into Poisson's eqn. (6.3), and, after integration, the electric field distribution is determined.

The n doped region can be solved similarly, substituting electrons with holes.

At each interface, between two adjacent segments, the continuity of the carrier concentration, the electric field and their derivatives are imposed. In this way it is possible to determine uniquely the carrier and electric field profiles over all the doped regions.

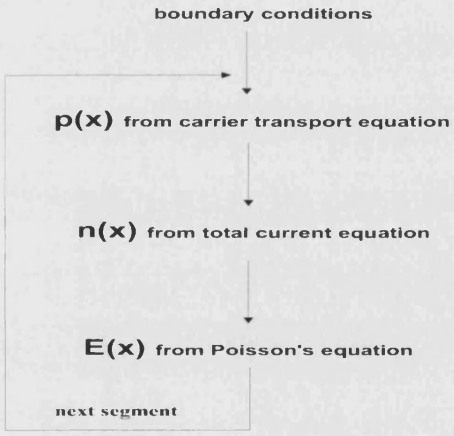


Figure E.1: Schematic flow-chart for solving the p doped region.

Appendix F

Voltage drop across the p-i-n junction

The excess carrier concentrations, Δp, Δn, inside the intrinsic region, near the junction edges, can be related to the equilibrium carrier concentrations, p_0, n_0, inside the heavily p, n, doped regions, respectively. The two junctions p-i and i-n can be considered as two different separated junctions, whose voltage drop across is V_{pi} and V_{in} respectively. From the p-n junction theory it is, with reference to Fig. [6.2],

$$\begin{aligned} p(-l/2^+) &\approx \Delta p(-l/2^+) = p_0 \exp \tfrac{V_{pi}}{V_t} \\ n(l/2^-) &\approx \Delta n(l/2^-) = n_0 \exp \tfrac{V_{in}}{V_t} \end{aligned} \tag{F.1}$$

Substituting eqns. (F.1) into the expression for the carrier distribution (6.9) it is

$$\begin{cases} p_0 \exp \frac{V_{pi}}{V_t} = C_1 \cosh \frac{-l/2}{L_i} + C_2 \sinh \frac{-l/2}{L_i} \\ n_0 \exp \frac{V_{in}}{V_t} = C_1 \cosh \frac{l/2}{L_i} + C_2 \sinh \frac{l/2}{L_i} \end{cases} \tag{F.2}$$

The constant C_1 and C_2 are defined as in (E.17), (E.18). L_i is the ambipolar diffusion length.

Multiplying both eqns. (F.2) it is

$$n_0 p_0 \exp \frac{V_{pi} + V_{in}}{V_t} = C_1^2 \cosh^2 \frac{l/2}{L_i} + C_2^2 \sinh^2 \frac{l/2}{L_i} \tag{F.3}$$

Defining the constants A, B as $A = C_1/J_0$, $B = C_2/J_0$, eqn. (F.3) becomes

$$n_0 p_0 \exp \frac{V_{pi} + V_{in}}{V_t} = {J_0}^2 (A^2 \cosh^2 \frac{d/2}{L} - B^2 \sinh^2 \frac{d/2}{L}) \tag{F.4}$$

The I-V characteristic, for the pairs of junctions, is then

$$J_0 = \frac{n_0 p_0}{\sqrt{A^2 \cosh^2 \frac{d/2}{L_i} - B^2 \sinh^2 \frac{d/2}{L_i}}} \exp \frac{V_{pi} + V_{in}}{2V_t} \tag{F.5}$$

While in the case of an ordinary p-n junction the relation between current and voltage is of the type $(\exp \frac{V_0}{V_t})$, with V_0 the applied bias across the junction, for a p-i-n junction with high

current injection, the slope is reduced by a factor of two.

In addition to the junction voltage drop, there is a voltage drop also across the intrinsic region, V_i: it is obtained by integrating the electric field profile over the spatial range of the intrinsic region.

Adding the two eqns. (5.5), (5.6) for the current, the electric field can be written as

$$E(x) = \frac{1}{e(\mu_p p + \mu_n n)}[J_0 - eV_t(\mu_n - \mu_p)\frac{dp}{dx}] \tag{F.6}$$

Integrating this expression inside the intrinsic region, is

$$V_i = -\int_{-l/2}^{l/2}(-E)dx = \frac{J_0}{e(\mu_p + \mu_n)}\int_{-l/2}^{l/2}\frac{dx}{p(x)} - \frac{V_t(\mu_n - \mu_p)}{(\mu_p + \mu_n)}\ln\frac{p(l/2)}{p(-l/2)} \tag{F.7}$$

As a result, the total voltage drop across the entire device is

$$V_{pin} = 2V_t \ln\left[\frac{J_0}{n_0 p_0}\sqrt{A^2\cosh^2\frac{l}{L_i} + B^2\sinh^2\frac{l}{L_i}}\right] + V_i \tag{F.8}$$

Appendix G

First order Perturbation Theory

An unperturbed system is in general described by the eigenvalue equation

$$U^0 E_n^0 = a_0 E_n^0 \tag{G.1}$$

where U^0 is the unperturbed differential operator, and a_0, E_n^0, the unperturbed eigenvalue and eigenmode, respectively.
If the operator U^0 is perturbed by a small amount, u, the corresponding eigenvalues and eigenfunctions, a_n, E_n, will be perturbed accordingly, [8]. They will satisfy the perturbed equation

$$U E_n = a_n E_n \tag{G.2}$$

U is the perturbed operator. The perturbed eigenvalues, a_n, and eigenfunctions, E_n, can be expanded in powers series, x^n, in proximity of their unperturbed value:

$$U = U^0 + ux \tag{G.3}$$

$$a_n = a_n^0 + a_n^1 x + a_n^2 x^2 + a_n^3 x^3 + \dots \tag{G.4}$$

$$E_n = E_n^0 + E_n^1 x + E_n^2 x^2 + E_n^3 x^3 + \dots \tag{G.5}$$

where the coefficients of the power terms with order higher than zero represent the corrections due to the perturbation.
Inserting eqns. (G.3), (G.4), (G.5) in (G.2) yields

$$(U^0+ux)\left(E_n^0+E_n^1x+E_n^2x^2+E_n^3x^3+\dots\right) = \left(a_n^0+a_n^1x+a_n^2x^2+a_n^3x^3+\dots\right)\left(E_n^0+E_n^1x+E_n^2x^2+E_n^3x^3+\dots\right) \tag{G.6}$$

Equating coefficients with the same power it is

$$\begin{array}{l|c} x^0: & U^0E_n^0 = a_n^0E_n^0 \qquad \text{same as (G.1)} \\ x^1: & U^0E_n^1 - a_n^0E_n^1 - a_n^1E_n^0 = -uE_n^0 \\ x^2: & U^0E_n^2 - a_n^0E_n^2 - a_n^2E_n^0 = -uE_n^2 + a_n^1E_n^1 \end{array} \tag{G.7}$$

If the perturbation, u, is small, only the first order correction term will be predominant, while the others can be neglected. Thus, only the first order equation is considered

$$U^0 E_n^1 - a_n^0 E_n^1 - a_n^1 E_n^0 = -u E_n^0 \tag{G.8}$$

The correction term, E_n^1, to the field profile, can be expanded in terms of the complete set, $\{E_m^0\}$, of mode solutions of the unperturbed eqn. (G.1)

$$E_n^1 = \sum_m c_{nm}^1 E_m^0 \tag{G.9}$$

Inserting expansion (G.9) into (G.8) and making use of eqn. (G.1) yields

$$\sum_m c_{nm}^1 \left(a_m^0 - a_n^0\right) E_m^0 - a_n^1 E_n^0 = -u E_n^0 \tag{G.10}$$

Multiplying eqn. (G.10) by the complex conjugate E_n^{0*} and integrating over the transverse spatial domain S, after some algebra, yields

$$\sum_m c_{nm}^1 \left(a_m^0 - a_n^0\right) \int_S E_m^0 E_n^{0*} \mathrm{d}x\mathrm{d}y - a_n^1 \int_S E_m^0 E_n^{0*} \mathrm{d}x\mathrm{d}y = -\int_S E_m^0 u E_n^{0*} \mathrm{d}x\mathrm{d}y \tag{G.11}$$

Because the eigenmodes $\{E_m^0\}$ are orthogonal it is

- for $m = n$

$$a_n^1 = \frac{\int_S E_n^0 u(x,y) E_n^{0*} \mathrm{d}x\mathrm{d}y}{\int_A \left|E_n^0\right|^2 \mathrm{d}x\mathrm{d}y} \tag{G.12}$$

- for $m \neq n$

$$c_{nm}^1 = \frac{\int_S E_m^0 u(x,y) E_n^{0*} \mathrm{d}x\mathrm{d}y}{\left(a_n^0 - a_m^0\right) \mathrm{d}x\mathrm{d}y} \tag{G.13}$$

Bibliography

[1] Robson P.N., Kendall P.C., "Rib Waveguide Theory by the Spectral Index Method", RSP, England, 1990

[2] Rozzi T., Mongiardo M., "Open Electromagnetic Waveguides", IEE, London, 1997

[3] Harrington R. F., "Time-Harmonic Electromagnetic Fields", McGraw-Hill, New York, 1961

[4] Kreyszig E., "Advanced Engineering Mathematics", John Wiley & Sons, New York, 1993

[5] Jeffreys H., Jeffreys B.S., "Methods of Mathematical Physics", CUP, Cambridge, 1972

[6] Sewell P., Benson T.M., Greedy S., Vukovic A., "Advances in the Spectral Index Method for OEIC Design", *International conf. On Transparent Optical Networks*, 9-11 June, Kielce, Poland, 1999, 137-138

[7] McKelvey J.P., "Solid State and Semiconductor Physics", Harper & Row, NY, 1966

[8] Fermi E., "Notes on Quantum Mechanics", Phoenix Science Series, The University of Chicago Press, Chicago, 1961

List of publications obtained from this thesis

- N. Cinosi, J. Sarma, "Computation of Injected Carrier Profile and consequent Optical Mode variation in Rib Waveguides", *Semiconductor and Integrated Optoelectronics Conference, SIOE'00*, Cardiff, 17-19 April 2000
- F. Causa, N. Cinosi, J. Sarma, "Comparison of Quasi-Analytic Methods to Compute the Modes of Optical Rib Waveguides", *Semiconductor and Integrated Optoelectronics Conference, SIOE'00*, Cardiff, 17-19 April 2000
- N. Cinosi, J. Sarma, "A Convenient Computation Scheme for the Evaluation of Optical Mode Control in Rib Waveguides and Couplers", *International Conference on Fiber Optics and Photonics, Photonics-2000*, 18-20 December 2000, Calcutta
- N. Cinosi, J. Sarma, "A Convenient Computation Scheme for the Evaluation of Optical Mode Control in Rib Waveguides and Couplers", *in Proc. SPIE. Vol. 4417. Photonics 2000: International Conference on Fiber Optics and Photonics*, 286-291
- N. Cinosi, J. Sarma, "Injected Carrier Profiles and Optical Mode Variations in Semiconductor Waveguides", *NKT Photonics Summer School in Advanced Photonics*, Copenhagen, August 2001
- N. Cinosi, J. Sarma, "A Quasi-Numerical Scheme for Modeling 2-Dimensional Carrier Profiles in Semiconductor Optical Waveguide Devices", *PREP 2002, Nottingham*, 17-19 April 2002
- N. Cinosi, J. Sarma, "Quasi-Numerical Solution of the 2-D Transport Equations for Carriers Injected into Semiconductor p-i-n junction Optical Waveguide Devices", *Semiconductor and Integrated Optoelectronics Conference, SIOE'02*, Cardiff, 25-27 March 2002
- N. Cinosi, J. Sarma, "Optical Propagation in Tapered Deep Ridge Waveguides", *International Conference on Optoelectronics, Fiber Optics and Photonics, Photonics-2002*, 14-16 December 2002, Mumbai
- N. Cinosi, J. Sarma, "Quasi Modal Field Propagation in Tapered Deep Ridge Dielectric Waveguides", *Semiconductor and Integrated Optoelectronics Conference, SIOE'03*, Cardiff, 14-16 April 2003
- N. Cinosi, J. Sarma, "An Iterative Method for Computing Carrier Distribution in 2-Dimensional Carrier Controlled Optical Guides", submitted to *IEEE Electronic Devices*

List of Figures

List of Tables